

AFOSR-TR- 78-1165

LEVEL

✓
A043 402

EXPERIMENTS IN TEXTURE PERCEPTION

2

AD A059630

DDC FILE COPY

Final Report

By

WHITMAN RICHARDS
DEPARTMENT OF PSYCHOLOGY

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASS. 02139

January 1978

Contract F44620-74-C-0076

DDC
RECEIVED
SEP 25 1978
RECEIVED

Prepared for
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
AIR FORCE SYSTEMS COMMAND
BOLLING AIR FORCE BASE, D.C. 20332

Handwritten signature and the letter 'D'.

78 07 20 166

Sponsor: Advanced Research Projects Agency ARPA Order 2765

AFOSR Project Monitors: A. Fregly and C. Hutchinson

ARPA Program Managers: C. Fields and A. Kibler

Approved for public release
distribution unlimited.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

<p>REPORT DOCUMENTATION PAGE</p> <p>AFOSR TR- 78-1165</p>		<p>READ INSTRUCTIONS BEFORE COMPLETING FORM</p> <p>2. GOVT ACCESSION NO</p> <p>3. RECIPIENT'S CATALOG NUMBER</p>	
<p>6. TITLE (and Subtitle)</p> <p>EXPERIMENTS IN TEXTURE PERCEPTION</p>		<p>9. PERFORMING ORG. REPORT NUMBER</p> <p>Final Report 1 Jun 74 through 30 Nov 77</p>	
<p>7. AUTHOR(s)</p> <p>10. Whitman A. Richards</p>		<p>8. CONTRACT OR GRANT NUMBER(s)</p> <p>15. 44620-74-C-0076, VV ARPA Order-2765</p>	
<p>9. PERFORMING ORGANIZATION NAME AND ADDRESS</p> <p>Massachusetts Institute of Technology 77 Massachusetts Avenue Cambridge, Massachusetts 02139</p>		<p>10. PROGRAM ELEMENT, PROJECT, AREA & UNIT NUMBERS</p> <p>61101E/P765</p>	
<p>11. CONTROLLING OFFICE NAME AND ADDRESS</p> <p>Advanced Research Projects Agency 1400 Wilson Boulevard Arlington, VA 22209</p>		<p>12. REPORT DATE</p> <p>11. January 1978</p>	
<p>14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)</p> <p>Air Force Office of Scientific Research (NL) Building 410 Bolling Air Force Base, D.C. 20332</p>		<p>15. SECURITY CLASS. (of this report)</p> <p>Unclassified</p>	
<p>16. DISTRIBUTION STATEMENT (of this Report)</p> <p>Approved for public release; distribution unlimited.</p> <p>12. 124 p.</p>			
<p>17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)</p>			
<p>18. SUPPLEMENTARY NOTES</p>			
<p>19. KEY WORDS (Continue on reverse side if necessary and identify by block number)</p> <p>Vision, Texture Perception, Graphics Display</p>			
<p>20. ABSTRACT (Continue on reverse side if necessary and identify by block number)</p> <p>Visual textures may be described completely by their spatial frequency components. For one-dimensional textures whose luminance varies only along the X-axis of the display, the descriptive elements are gratings that have sinusoidal modulations of luminance. Although any arbitrary one-dimensional "blurred" texture may require a very large number of sinusoidal components for its complete physical description, only four components are needed to create a texture</p>			

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

403 698

elt

**Best
Available
Copy**

Unclassified

-116-

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

that appears equivalent to the human observer. Thus, the human visual system does not act like a spectral analyzer, but rather appears to process spatial frequency information by filtering operations similar to that performed in color vision. In the more general case, textures will have luminance distributions varying in two dimensions (i.e. in both X and Y). If a two-dimensional texture is created with orthogonal luminance profiles (whereby the axis orientations of X and Y are 90° apart), then the X and Y profiles are independent and four spatial frequencies will be needed for X and four for Y. The most general case of texture equivalence, where many orientations are present in a texture, has not yet been solved. However, preliminary experiments (by M. Riley) show that orientation equivalence can be attained by utilizing only four independent orientations. This constraint suggests an upper bound of sixteen on the number of fixed spatial frequencies required to create an equivalence to any two-dimensional texture pattern. However, where control over the maximum visual angle can be maintained, twelve spatial frequencies may suffice for practical purposes, especially if the basic waveform of the primary components can be pre-programmed. These limitations of human visual processing suggest that data transmission rates of textural information (such as homogeneous surfaces) can be greatly compressed.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

3

Summary

Visual textures may be described completely by their spatial frequency components. For one-dimensional textures whose luminance varies only along the X-axis of the display, the descriptive elements are gratings that have sinusoidal modulations of luminance. Although any arbitrary one-dimensional "blurred" texture may require a very large number of sinusoidal components for its complete physical description, only four components are needed to create a texture that appears equivalent to the human observer. Thus, the human visual system does not act like a spectral analyzer, but rather appears to process spatial frequency information by filtering operations similar to that performed in color vision. In the more general case, textures will have luminance distributions varying in two dimensions (i.e. in both X and Y). If a two-dimensional texture is created with orthogonal luminance profiles (whereby the axis orientations of X and Y are 90° apart), then the X and Y profiles are independent and four spatial frequencies will be needed for X and four for Y. The most general case of texture equivalence, where many orientations are present in a texture, has not yet been solved. However, preliminary experiments (by M. Riley) show that orientation equivalence can be attained by utilizing only four independent orientations. This constraint suggests an upper bound of sixteen on the number of fixed spatial frequencies required to create an equivalence to any two-dimensional texture pattern. However, where control over the maximum visual angle can be maintained, twelve spatial frequencies may suffice for practical purposes, especially if the basic waveform of the primary components can be pre-programmed. These limitations of human visual processing suggest that data transmission rates of textural information (such as homogeneous surfaces) can be greatly compressed.

ACCESSION for	
DTIC	White Section <input checked="" type="checkbox"/>
DDI	Buff Section <input type="checkbox"/>
UNANNOUNCED	<input type="checkbox"/>
JUSTIFICATION	
BY	
DISTRIBUTION/AVAILABILITY CODES	
Dist.	AVAIL. and/or SPECIAL
A	



Figure 1.1

The special effects graphics display used to drive two TV monitors with 440 X 440 X 64 resolution. The device has 64K of 18 bit refresh memory, reprogrammable for use as PDP 11 core. The disk capacity is 2.5 million words. See Appendix for complete description of system.

ACKNOWLEDGEMENT

E.H. Black guaranteed the success of this project by his contributions to the design and development of the "Texture Machine," whose capabilities have still not been fully explored. This contribution was critical to the project. Others who generously contributed to these texture studies were L. Dubose (Hycor), H. Lieberman, D. Marr, H.K. Nishihara, S. Purks, A. Polit, R.F. Quick, M. Riley, R. Spitzberg, B. Tressler (Hycor), A. Vezza and A. Witkin.

CONTENTS

	<u>Page</u>
Summary	i
Acknowledgement	iii
Index to Figures	vi
I. Introduction	1
II. One-Dimensional Texture Perception	3
1. Definition	
2. Method	
3. Linearity Assumption	
4. Preliminary Selection of Primaries	
5. Results	10
i) Free Viewing	
ii) Additivity Test	
iii) No Eye Movements--Central Fields	
iv) No Eye Movements--Extra-foveal Fields	
v) Oblique Texture Matches	
vi) Square-Wave Primaries	
vii) Linearity	
viii) Reduced Texture Space	
III. Two-Dimensional Texture Matches	37
1. Basic Contrast Sensitivity Data	
2. Texture Matches	
a) Sums	
b) Products	
IV. Random-Dot Textures	43
N-Gram Patterns	
Method of Generating Textures	
Illustration of Generation Method	
Test Procedures	
Results	
Nth Order Patterns	

CONTENTS (Cont.)

	<u>Page</u>
V. Site of Texture Matching	58
i) Binocular Matching	
ii) Adaptation	
iii) Estimate of Physiologic Primaries	
VI. Generalized Colorimetry	67
1. Principal Features of Colorimetric Approach	
2. Generalized Complements	
3. Examples of Generalized Complement Technique	
Texture Metamers	
Orientation Discrimination	
Flicker Discrimination	
4. Derivation of Matching Functions	
5. Examples of Derivation of Response Functions	
i) Spatial Frequency Matches	
ii) Temporal Frequency Matches	
6. Intensive Versus Extensive Variables	
7. Hierarchical Processing	
VII. Major Conclusions	86
VIII. Significance	88
IX. Bibliography	89
X. Appendices	98
DD143	115

78 07 20 166

Index to Figures

<u>Fig. No.</u>		<u>Page</u>
1.1	Texture Machine	ii
2.1	Examples of one-dimensional textures	6
2.2	Matching a "white noise" texture	6
2.3	Method for the preliminary selection of primaries	8
2.4	Example of texture match at 2.2 c/deg made with free viewing . . .	11
2.5	Texture matching functions	14
2.6	Example of texture match at 1.5 c/deg from Table II	16
2.7	Example of texture match at 8.1 c/deg from Table II	18
2.8	Texture match with square waves at 0.9 c/deg from Table IV	22
2.9	Matches between 64 and 3 gray levels	24
2.10	Further matches between 3 and 64 gray levels	25
2.11	Statistical and "Retinex" description of textures	27
2.12	Comparison of bars, checks and hexagons	29
2.13	Reduced texture space	36
3.1	Two-dimensional sensitivity functions	39
3.2	Two-dimensional patterns	42
4.1	N-gram patterns	47
4.2	Additional N-gram patterns	50
4.3	A discriminable 3rd order pattern	54
5.1	Effect of grating adaptation on texture matches	61
6.1	Method of Generalized Colorimetry	70
6.2	Match to checks with randomly distributed gray levels	76
6.3	Matches to lines with random orientations	78
6.4	Flicker matching functions	82

Experiments in Texture Perception

I. Introduction

Texture, like color, is one of the primary properties of an object (Metzger, 1926; Koffka, 1935). Yet our knowledge of the texture recognition process of the human observer is meagre. Previous studies of texture may be crudely divided into three categories:

- 1.) texture gradients as shown and their roles in slant and depth perception (Gibson, 1950; Gruber and Clark, 1956; Wohlwill, 1962; Flock and Moscatelli, 1964; Kraft and Winnick, 1967);
- 2.) texture discrimination and its relation to the statistical properties of the display (Jones and Higgins, 1947; McBride and Reed, 1952; Green et al, 1959; Stultz and Zweig, 1959; Julesz, 1962, 1965; Pickett, 1962, 1964, 1967) and
- 3.) the search for continua suitable for an objective definition of "texture" (Jones and Higgins, 1945; Rosenfeld, 1967; Pickett, 1968; Minsky and Papert, 1969; Julesz, 1971).

Although clearly relevant to these previous studies, our primary approach to texture perception is entirely new and falls into still another category. The novelty of the new approach is that it is concerned only with describing textures that appear equivalent to the observer, rather than trying to specify the physical characteristic that will differentiate between all textures. The attempt to describe equivalent textures is analogous to the development of color science where the primary concern was to identify spectral compositions that would appear equivalent to the human observer. Such energy distributions that were physically different but appeared equivalent were called metamers. Our approach to texture perception is to describe such metamers.

The first step in describing color equivalences was the recognition of the dimension of wavelength into which the visual scientist could map the components of all spectral lights. Texture may also be described in exactly the same way except the relevant dimension is now spatial frequency (DePalma and Lowry, 1962; Robson, 1966; Bryngdahl, 1966; Campbell and Robson, 1968). Thus, if at the onset only one-dimensional textures are considered, then Fourier's theorem states that any such texture may be adequately described by the magnitude of its sinusoidal components. These components are of course merely sine-wave gratings which when added together in suitable proportions will physically recreate the one-dimensional texture pattern. Thus, the dimension of spatial frequency can be used to describe all possible one-dimensional textures in exactly the same manner that chromatic wavelength is used to describe all possible colors.

In color it was discovered during the last century that only three suitably chosen wavelengths were needed to generate equivalences to all possible physically realizable colors (Maxwell, 1855; Wright, 1928; Guild, 1931). This property of color equivalences is imposed by the fact that human color perception is based upon the energy passed through only three independent filters each having a different wavelength characteristic (Stiles and Burch, 1959; Brown and Wald, 1964; Marks, Dobelle and MacNichol, 1964). This report shows that texture perception follows a similar principle: namely that all one-dimensional textures can be suitably matched by only a small number of suitably chosen sine-wave gratings put together in the right proportions.

II. One-Dimensional Texture Perception

1. Definition

For purposes of texture matching, texture is defined as an attribute of a field having no components that appear enumerable. Furthermore, the field should appear relatively homogeneous without obvious gradients. The intent of this definition is to require the observer to attend to the global properties of the display--i.e., its "coarseness," "bumpiness" or "fineness"--rather than analyzing the pattern segment by segment. Physically, such (aperiodic) patterns that are not enumerable are generated by a stochastic as opposed to a deterministic process (Pickett, 1968). Perceptually, however, the set of all patterns without obvious enumerable components will include many deterministic (and even periodic) textures. Because our criterion for enumerability was subjective rather than objective, many of our patterns actually contained repetitive elements which were not obvious but were occasionally noticed only upon refined inspection. To further minimize enumerable, periodic components of the patterns, all displays contained only frequency components that were prime ratios to one another.

2. Method

Complex patterns containing any finite number of sine-wave components were generated using a Special Graphics System controlled by a PDP 11/10 computer (see Appendix II and III). The display consisted of two independently controlled 14" video monitors with 440 x 440 element resolution and a 64 level gray scale with a P4 white phosphor. The mean luminance level was 20 cd/m². The refresh cycle was 32 msec, due to interlacing of the horizontal raster.

For most of the experiments that follow, the subject sat 200 cm from the TV monitors. In his lap, he held a control box that allowed him to adjust on-line the contrasts of up to six sinusoidal components of the displayed pattern. For threshold measurements, of course, only one knob needed to be adjusted. For texture matches, however, up to, but no more than four knobs needed adjustment for any given trial. This limitation, as will be seen shortly, was an empirical finding reflecting a limitation in human visual processing, and was not a limitation imposed by the equipment. Generally for these texture matches, either one or two of the left-hand knobs controlled components of the left screen (or left panel of the display), whereas the remaining three knobs controlled the right-hand components of the right screen (or right panel when the two matching fields were near-adjacent).

For all texture matches, the task of the subject was to adjust the contrast of the sinusoidal components of the pattern in both the left and right fields so that both fields (or panels) looked equivalent. In addition to the more formal definition of equivalence given earlier, we also used a more intuitive description of equivalence: "Make both panels look like they had been cut out from different regions of the same rug." Or, alternately, "Can one texture be considered an extension of the other?". It was also necessary to stress that equivalence did not mean physical identity whereby the phases and number of cycles matched exactly in each panel. If the textures in the two panels were judged not to be equivalent, then the contrasts of the components of one or both textures were altered by the subject until the best texture match was obtained.

These matches were then ranked by the subject on a scale from poor, fair, good, very good, and excellent, with the latter category implying physical indistinguishability. Over 90% of our final results are based on "very good" or better ratings.

3. Linearity Assumption

Clearly it is an impractical task to match all possible textures with a few, fixed spatial frequencies. The number of such textures is just too large. To circumvent this obstacle, a linearity assumption is made:

Any (subjective) texture may be characterized by the linear superposition of the Fourier components of the actual pattern.

Although this assumption regarding the behavior of the visual system is known to be false, particularly at high contrasts (Davidson, 1968; Cornsweet, 1970; Franzen and Berkley, 1975), the approximation is good at low contrasts (Campbell and Robson, 1968; Henning et al, 1975; Abadi and Kulikowski, 1973; Kulikowski, 1976; Quick and Reichart, 1975; Graham, 1977). With this approximation it is then necessary only to specify an equivalence between each pure sine-wave pattern and the chosen fixed spatial frequencies (primaries) in order to specify matches to all possible textures. The procedure is thus directly analogous to that used to specify color matches in colorimetry (Wyszecki and Stiles, 1967). And, like colorimetry one of the primaries will always be added as a "desaturant" to the test frequency, with the combination to be matched by the remaining two primaries. When a primary is added as a "desaturant" to the test frequency, the primary will assume negative values.

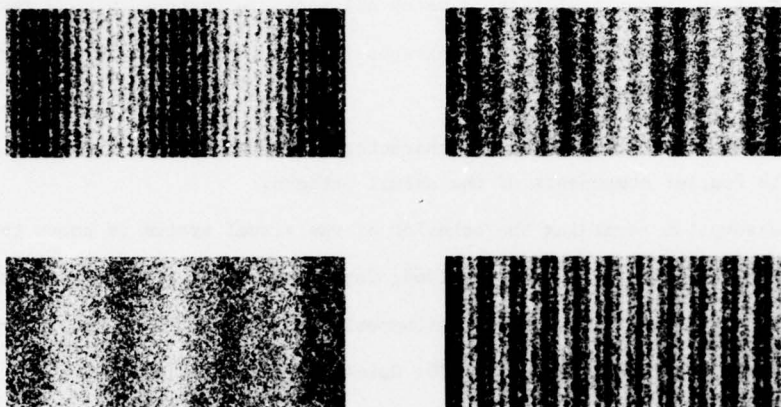


Figure 2.1

Four examples of one-dimensional textures composed of only a few sinusoidal components. As the number of components increases, the textures approach those shown below.



Figure 2.2

The pattern to the left contains noise restricted to the range 0.2-20 c/deg when viewed at 50 cm. The texture on the right, which is considered a texture metamer, contains only three frequency components, 0.53, 2.4 and 6.5 c/deg.

4. Preliminary Selection of Primaries

Our first task was to determine just how closely spaced the spatial frequency spectrum from 0.2 to 30 c/deg should be sampled. In order to set a lower bound on the number of fixed primary spatial frequencies needed, texture matches were first made to "white noise" patterns that contained a random selection of sinusoidal frequencies and amplitudes. It was found that three spatial frequencies were sufficient to make such matches, as illustrated in Fig. 2.2. This solution also set an upper bound of six on the number of primaries needed (see Generalized Colorimetry section).

To determine the location and exact number of the primaries required, a fixed form of a texture primary was assumed, as characterized by the inset to Fig. 2.3. Along a log spatial frequency axis, the primary function has one positive lobe flanked by two negative lobes. (Both lobes have been found to be necessary to create texture metamers.) We can now ask the experimental question of how large a separation may be present between the location of adjacent primaries for texture equivalence to hold. The answer is obtained by measuring the acceptability of texture matches between frequency f (a variable) and the primaries which bear a fixed relation to f . The relation is as follows:

$$0.5 (f) + A (k^{3/2}f) + B (k^{-3/2}f) \equiv C (k^{1/2}f) + D (k^{-1/2}f) \quad (1)$$

where the contrast of f is held fixed at 0.5 and $A - D$ are the measured contrasts of the primary frequencies ($k^x f$).

Fig. 2.3 shows the values of the coefficients $A - D$ for values of f ranging from $1/4$ to 20 c/deg. These values do not change much as k is altered from 2 to 3, but the acceptability of the texture matches does. If free eye movements and viewing are allowed, excellent texture equivalences can be obtained only if k is less than 2.4. Thus, the "half-width" of a primary is of this magnitude, and four primaries can span a range of only $2.4^4 = 31$. For practical purposes, however, this range is quite acceptable, covering all patterns except those with luminance "gradients" less than 25% per degree.

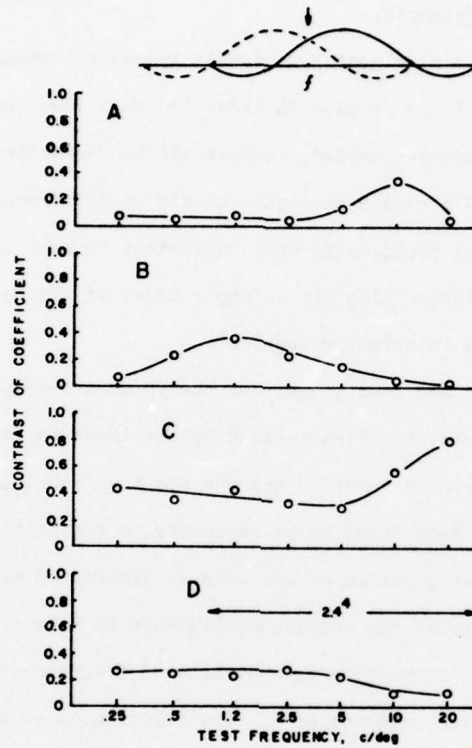


Figure 2.3

Test to determine the minimum bandwidth necessary for texture primaries. The waveform of the primaries is shown in the inset. See text equation (1) for the description of the relations between the primaries. Each graph shows the contrast of a primary needed to match the spatial frequencies given on the abscissa.

Note that all coefficients have constant values over a wide portion of the range examined. This important property permits a further simplification, for if the coefficient values were flat everywhere, then texture matches would be invariant over visual angle or fixation distance. At the lower spatial frequencies, a partial size constancy is obtained. At higher spatial frequencies, the failure in constancy is due to failures in the resolution of the highest spatial frequency components.

The above constraints together with further pilot studies then led to the following choice of primaries under free viewing conditions: 11, 6.3, 3.2, 1.5, .9, .3 c/deg.

Without eye movements, the above set could be reduced to: 11, 6.3, 3.2, .9 c/deg.

5. Results

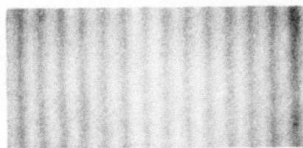
i) Free Viewing

Whether or not eye movements are allowed makes a big difference with regard to the spacing of the primary spatial frequencies and their contrasts. As the eyes move across the pattern, both spatial and temporal cues are present. The temporal cues are particularly important for enhancing low spatial frequencies (DeLange, 1958; Kelly, 1977; Koenderink, 1972), and consequently more primaries are needed in this region of the spatial frequency spectrum.

Table I gives the average one-dimensional texture matching functions obtained from four observers using free eye movements and viewing two 7° wide by 6° high fields seen at 200 cm. The six fixed primaries are sufficient over the range examined to create "very good" to "excellent" matches, provided that the texture components are sinusoidal, and provided that desaturation is allowed (as indicated by the negative contrasts). Figure 2.4 illustrates a match made to 2.2 c/deg.

These values in Table I can be used to predict texture equivalences in a manner similar to the use of distribution functions in colorimetry to predict the equivalence between different spectral lights (see Wysecki and Stiles, 1967, for use of distribution functions). However, because of the non-linear behavior of the visual system in the neighborhood of sharp edges (Cornsweet, 1970), only textures that appear "fuzzy" or "blurred" can be matched using the functions listed in Table I. For textures containing a significant number of sharp edges or lines, square-wave primaries must be used. An appropriate set of these functions will be described later.

TEST



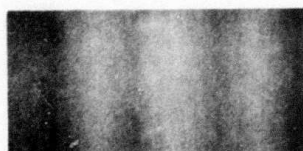
0.5 (2.2)

+



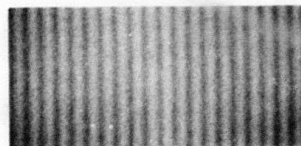
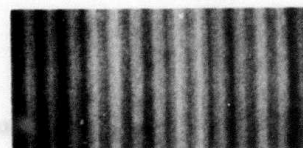
0.09 (6.3)

+



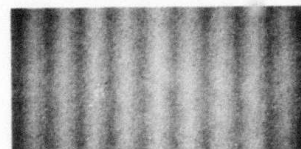
0.22 (.57)

||



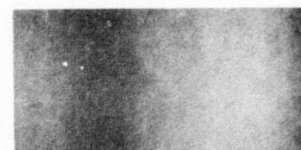
0.38 (3.2)

+



0.30 (1.5)

+



0.07 (.16)

||

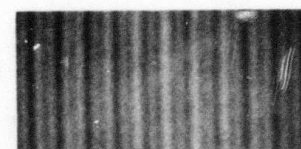


Figure 2.4

Texture match to 2.2 c/deg made with free viewing. Each contrast used is followed by its spatial frequency given in parentheses. Values are selected from Table I. Proper viewing distance is 50 cm.

TABLE I

TEXTURE MATCHES TO VERTICAL SINUSOIDS
(7x6 deg field--free viewing)

TEST FREQUENCY	PRIMARY FREQUENCY, c/deg						THRESHOLD CONTRAST	
c/deg	10.7	6.3	3.2	1.5	.57	.16	Measured	Predicted
31.	+.08(.04)	-.01(.01)	+.01(.01)	-.01(.01)			.86(.09)	.68*
22.5	+.18(.02)	-.01(.02)	+.02(.01)	-.02(.01)			.37(.11)	.24*
15.5	+.30(.03)	-.11(.01)	+.06(.03)	-.02(.01)			.153(.03)	.19*
10.7	<u>.50</u>						.075(.02)	.091
9.5	+.60(.07)	+.02(.02)	-.01(.01)	0			.070(.02)	.074
8.1	+.54(.07)	+.18(.07)	-.10(.05)	+.01(.01)			.056(.01)	.076*
7.0	+.12(.05)	+.47(.04)	-.06(.03)	+.02(.02)			.066(.01)	.055
6.3		<u>.50</u>					.053(.01)	.058
5.4	-.02(.03)	+.52(.04)	+.03(.04)	-.02(.03)	+.01(.02)		.040(.01)	.049*
4.5	-.13(.05)	+.54(.08)	+.17(.01)	-.10(.04)	+.04(.05)		.035(.01)	.041
3.8	-.01(.01)	+.06(.05)	+.51(.03)	-.08(.04)	+.08(.05)		.033(.01)	.032
3.2			<u>.50</u>				.033(.01)	.032
2.7		-.03(.02)	+.47(.03)	+.08(.07)	-.04(.05)	+0	.034(.01)	.032
2.2		-.09(.04)	+.38(.04)	+.30(.07)	-.22(.05)	+.07(.06)	.033(.02)	.032
1.8		-.05(.04)	+.17(.12)	+.18(.04)	-.15(.05)	+.07(.01)	.037(.02)	.033
1.5				<u>.50</u>			.039(.01)	.039
1.2		+.05(.04)	-.10(.03)	+.48(.05)	+.25(.05)	-.14(.08)	.038(.01)	.040
.93		+.05(.03)	-.11(.04)	+.42(.06)	+.31(.04)	-.26(.10)	.046(.02)	.047
.70		+.02(.02)	-.04(.02)	+.16(.06)	+.54(.02)	-.22(.04)	.052(.02)	.057
.57				<u>.50</u>			.079(.02)	.079
.41			+.02(.01)	-.04(.03)	+.37(.06)	+.45(.13)	.089(.04)	.076
.30			+.01(.01)	-.04(.03)	+.18(.06)	+.53(.06)	.106(.04)	.112
.22			+.01(.01)	-.02(.02)	+.09(.05)	+.51(.09)	.151(.07)	.134
.16						<u>.50</u>	.174(.10)	.173

N = 4

(Values in parentheses indicate one third the range of the observers' settings.)

ii) Additivity Test

The last column of Table I gives the threshold contrast sensitivity predicted by the texture matching functions. These values are calculated by determining the weights of each primary needed to match the reciprocal of the contrast threshold of that primary. This weight for each primary is then applied to all its distribution values (i.e. to the entire column of matching contrasts given for that primary). The weighted contrasts for each column are then summed together. The reciprocal of the weighted sum should equal the contrast sensitivity for each test frequency. As can be seen by comparing the last column of Table I with the measured thresholds given in the next to last column, the agreement is within 20%--a value consistent with experimental error. Thus, the only serious departure from additivity is at the highest spatial frequency.

iii) No Eye Movements--Central Fields

If no eye movements are allowed, then temporal information about the texture pattern is lost. Under these conditions, low spatial frequency sensitivity is impaired (especially because large eye movements are required to produce a significant contrast change), and the number of primaries required is reduced from six to four. Table II gives distribution functions for four sinusoidal primaries that cover the range from $1/6$ to 30 c/deg. These functions apply to a field 3 deg wide by 2° high, centered $1\frac{5}{6}$ deg off the fovea. (The decentering is the result of a $2/3$ deg separation between the two texture panels, with fixation midway between each.) The general form of these functions can be seen more clearly by inspecting Figure 2.5.

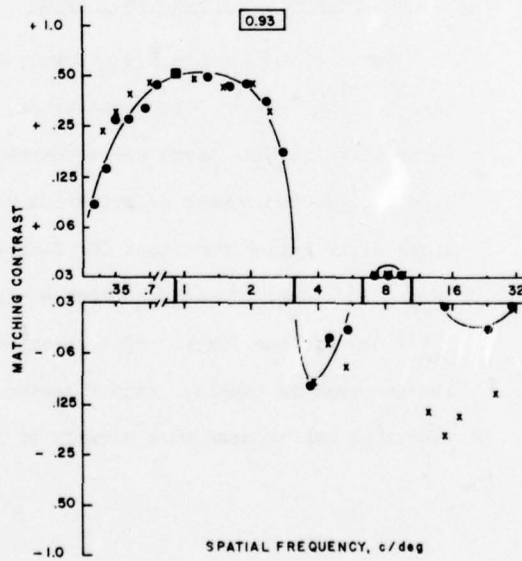
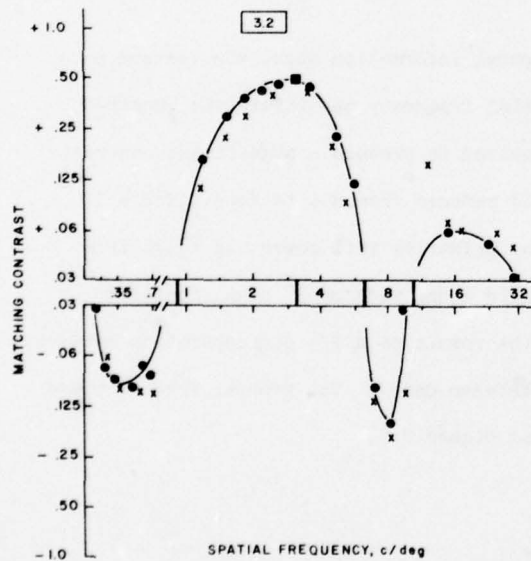
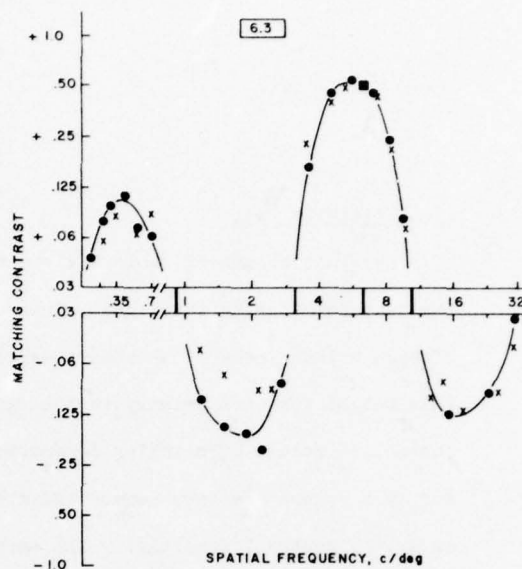
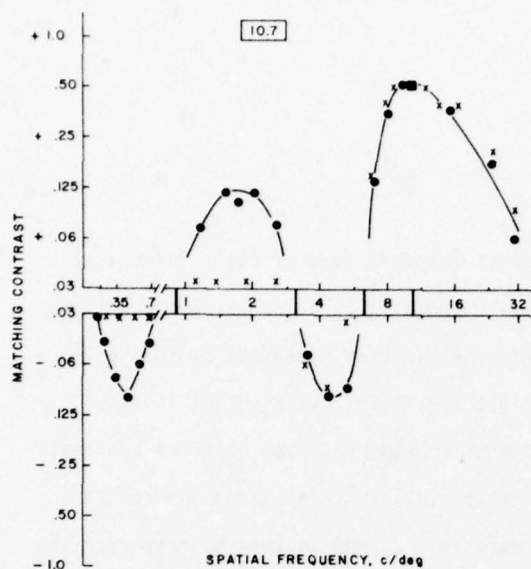


Figure 2.5

Texture Matching functions averaged over 5 observers, using vertical sinusoid grating test patterns of contrast 0.5 and no eye movements. Crosses are matches to textures oriented at 45° for WR, with frequency scale adjusted by a factor of 1.6x.

TABLE II

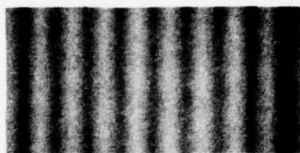
TEXTURE MATCHES TO VERTICAL SINUSOIDS
(3x2 deg field--no eye movements)

TEST FREQUENCY	PRIMARY FREQUENCY, c/deg				THRESHOLD CONTRAST	
	c/deg	10.7	6.3	3.2	0.93	Measured Predicted
31		+.06(.04)	-.01(.02)	+.01(.01)	0(0)	.91(.05) .94
22.5		+.24(.03)	-.09(.03)	+.05(.02)	-.04(.03)	.44(.13) .49
15.5		+.37(.05)	-.12(.04)	+.06(.04)	-.01(.02)	.20(.10) .24
10.7		<u>.50</u>				.115(.05) .115
9.5		+.52(.06)	+.08(.03)	-.03(.02)	0(0)	.107(.05) .093
8.1		+.37(.11)	+.24(.13)	-.15(.07)	+.02(.04)	.079(.03) .098
7.0		+.13(.06)	+.44(.02)	-.10(.03)	+.02(.03)	.068(.04) .068
6.3			<u>.50</u>			.053(.03) .057
5.4		-.09(.07)	+.53(.06)	+.12(.04)	-.04(.03)	.052(.02) .049
4.5		-.10(.05)	+.43(.03)	+.21(.07)	-.05(.03)	.052(.03) .052
3.8		-.05(.04)	+.17(.10)	+.44(.04)	-.09(.07)	.057(.03) .055
3.2				<u>.50</u>		.055(.03) .058
2.7		+.07(.04)	-.08(.03)	+.44(.07)	+.18(.11)	.052(.03) .058
2.2		+.12(.07)	-.19(.05)	+.41(.05)	+.37(.06)	.051(.02) .058
1.8		+.10(.04)	-.17(.03)	+.39(.05)	+.45(.07)	.053(.02) .054
1.5		+.12(.06)	-.14(.05)	+.29(.10)	+.42(.07)	.069(.03) .063
1.2		+.07(.06)	-.10(.05)	+.17(.08)	+.50(.07)	.082(.03) .072
.093					<u>.50</u>	.096(.03) .096
.66		-.05(.02)	+.06(.04)	-.08(.06)	+.42(.04)	.114(.03) .138*
.57		-.06(.03)	+.07(.03)	-.07(.05)	+.32(.05)	.15(.03) .18
.41		-.10(.01)	+.12(.03)	-.10(.04)	+.27(.07)	.23(.09) .21
.30		-.07(.02)	+.09(.03)	-.08(.03)	+.27(.07)	.39(.15) .21 "
.22		-.04(.03)	+.07(.03)	-.07(.03)	+.13(.07)	.41(.12) .48
.16		-.02(.02)	+.04(.03)	-.03(.03)	+.08(.05)	.56(.16) .59

N = 5

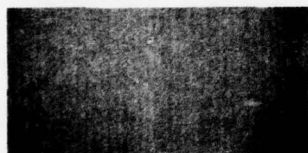
(Values in parentheses indicate one third the range of the observers' settings.)

TEST



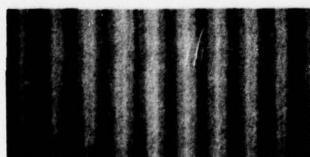
0.5 (1.5)

+



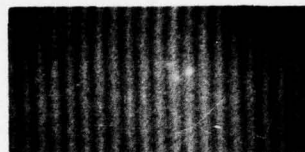
0.14 (6.3)

||



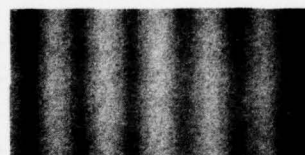
0.12 (10.7)

+



0.29 (3.2)

+



0.42 (.93)

||

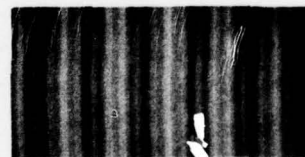


Figure 2.6

Texture match to 1.5 c/deg made with fixation between the lower pair of textures. Each contrast used is followed by its spatial frequency. Values are selected from Table II. Proper viewing distance is 50 cm.

These functions of Table II are more representative of the true spatial filtering properties of central human vision than are the previous functions listed in Table I. With eye movements minimized, the data of Table II are not as confounded by temporal and motion cues that help the observer enumerate the components of the textures. However, like Table I, these data apply only to "blurred" or "fuzzy" textures that are typical of patterns constructed from a small number of sinusoids having random phase relations. Figures 2.6 and 2.7 illustrate two texture matches based on the Table II values. (Fixation should be held midway between the two composite gratings.)

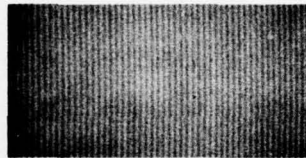
The last two columns of Table II compare again the contrast thresholds for the test frequencies predicted from the distribution functions. Once again, the agreement between the predicted and measured values is good, except at two very low spatial frequencies. Thus, over a wide range of spatial frequencies, "additivity" holds, demonstrating a linear property of the distribution functions.

iv) No Eye Movements--Extra-foveal Fields

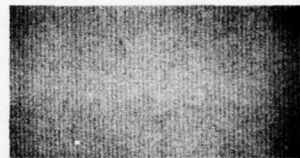
Because of equipment limitations imposed by screen size and raster resolution, central and peripheral texture equivalences were examined separately. To obtain texture matching functions for more eccentric retinal positions, a 7° wide by 2° high field was created on each monitor, and fixation was held between the two fields, which were separated by 6 deg. Thus, these fields merely extended the spatial range of the panels used to obtain the data of Table II.

Contrary to expectation, it was not necessary to change the spatial frequency primaries for these more peripheral matches. Thus, the primaries of Table II and III, which summarize the peripheral texture matches, are the same. This important result suggests that the same four spatial frequency filters underly human texture analysis in

TEST



0.5 (8.1)



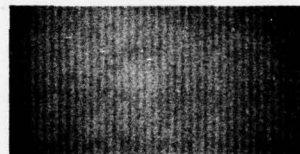
0.37 (10.7)

+



0.15 (3.2)

"



0.24 (6.3)

+



0.02 (1.3)

"

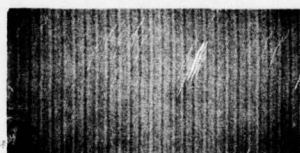


Figure 2.7

Texture match to 8.1 c/deg made with fixation held between the lower pair of textures.
Values chosen from Table II.

TABLE III
TEXTURE MATCHES TO VERTICAL SINUSOIDS
(7x2 deg field--no eye movements)

TEST FREQUENCY	PRIMARY FREQUENCY, c/deg				THRESHOLD CONTRAST
c/deg	10.7	6.3	3.2	.93	
31	0	0	0	0	1.0
22.5	+.16	0	0	0	.80
15.5	+.26	-0	+0	-0	.45
10.7	<u>.50</u>				.25
9.5	+.48	+.54	-.10	0	.22
8.1	+.15	+.39	-.14	+.03	.18
7.0	+.10	+.48	-.08	0	.13
6.3		<u>.50</u>			.11
5.4	-.22	+.54	+.11	-.01	.10
4.5	-.12	+.21	+.40	-.01	.075
3.8	-.03	+.03	+.50	-.03	.072
3.2			<u>.50</u>		.070
2.7	0	-.15	+.43	+.31	.065
2.2	0	-.21	+.28	+.45	.063
1.8	0	-.18	+.21	+.45	.067
1.5	0	-.10	+.18	+.43	.071
1.2	0	-.04	+.10	+.47	.079
.93				<u>.50</u>	.090
.66	0	+.04	-.08	.37	.097
.57	-.15	+.12	-.13	.38	.16
.41	-.08	+.09	-.09	.28	.20
.30	-.09	+.08	-.09	+.19	.37
.22	-.06	+.06	-.10	+.19	.55
.16	-.09	+.09	-.09	+.14	.52

N = 2

the central ten degrees of vision. Although the sensitivities of these four "channels" may change with retinal eccentricity, as indicated by the slightly different distribution functions, their bandpass characteristics do not. This result suggests that any scaling (magnification) of the location of spatial frequencies "channels" with eccentricity is inappropriate for suprathreshold pattern recognition (Hilz and Cavonius, 1974; Wilson and Giese, 1977; Spekreijse and van der Tweel, 1977; Limb and Rubinstein, 1977; Cowan, 1977). Thus, texture processing probably follows the same guidelines as color processing: the channels remain the same with retinal eccentricity although relative sensitivities are altered.

v) Oblique Texture Matches

Texture matches were also made by two subjects with the patterns seen at 45 deg by tilting one's head. Two 3x2 degree panels were used, as for the results of Table II. In addition, all subjects contributing to the Table II distribution functions were asked to grade their texture matches with head tilted at 45 deg. In general, the quality of most matches improved when viewed at 45°, suggesting a poorer resolution of oblique patterns. Such a conclusion would be premature, however.

In the region for test frequencies between 1.2 and 2.2 c/deg, all subjects consistently reported that the matches made at 90° orientation (vertical) became worse if viewed at 45° orientation. This decrement in quality implies that the 90° primaries are unsuitable for 45° matches in this region of test frequencies. Preliminary exploration was then initiated to discover a new set of primaries that would yield "very good" to "excellent" texture matches over the entire range from 1/6 to 30 c/deg. The best set of primaries found to date is merely the 90° set scaled to lower spatial frequencies by a factor of 0.7. Once this scaling factor is applied, then the distribution functions resemble those of Table II. (Specifically, viewing the 10.7 c/deg primary becomes

7.6 c/deg, etc., and all test frequencies are reduced by 0.7x). The plusses in Figure 2.5 are the 45° results appropriately scaled so they can be superimposed upon the averaged 90° data.

vi) Square-Wave Primaries

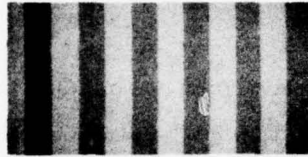
Many textures consist of patterns that have a large number of sharp lines or edges. Such textures cannot be adequately matched using only four sine-wave primaries, for additional harmonics must be included to create the edge effects. To match textures with lines or edges, the primary basis must be changed to a square wave-form.

Four suitable square-wave primaries are clustered together in the frequency range of 3.0 to 12 c/deg. Lower primary frequencies are not suitable, whereas spatial frequencies higher than 12 are inefficient. Table IV gives matching functions for the primaries of 3.1, 4.8, 7.0 and 10.7 c/deg, using square-wave test frequencies and for matches made without eye movements. The nature of these matches is such that the final match for all test patterns looks very similar--like a "white" noise texture. (Figure 2.8 shows a sample match to 0.9 c/deg.) Because of this desaturating effect of the mixtures, and because of masking of smooth gradients by edges, the square-wave primaries can also be used to describe equivalent textures for sinusoidal test frequencies.

One limiting case using square-wave luminance profiles is when texture patterns are created from narrow lines of equal width, but varying gray level. Fig. 2.9 shows such a pattern where the left portion of the figure has 64 gray levels randomly assigned to each bar or stripe. The other half of the pattern is made up of lines of the same width, but the gray level of each line is chosen from only three gray levels. Figure 2.10 shows patterns constructed in a similar manner, but using different square-wave spatial frequencies (parentheses) and contrasts. Note that suitable texture

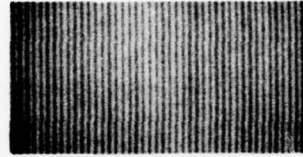
31

TEST



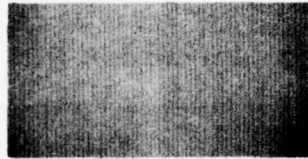
0.5 (0.9)

+



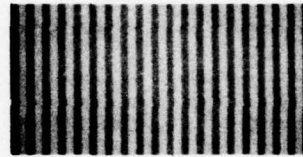
0.57 (7.0)

+



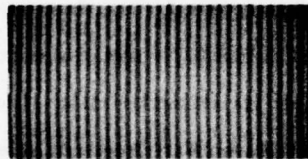
0.32 (10.7)

+



0.67 (3.1)

"



0.64 (4.8)

"

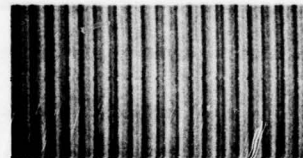
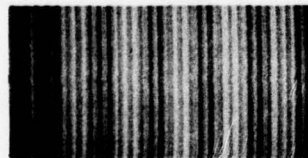


Figure 2.8

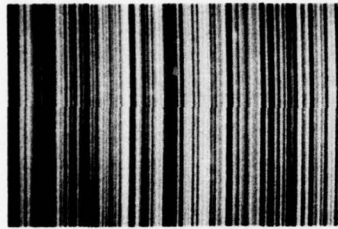
Texture match to a 0.9 c/deg square-wave made with fixation between the lower pair of textures. Values are chosen from Table IV. Proper viewing distance is 50 cm.

TABLE IV
TEXTURE MATCHES USING SQUARE-WAVE GRATINGS
(3x2 deg field--no eye movements)

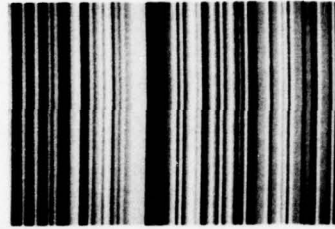
TEST FREQUENCY	PRIMARY FREQUENCY, c/deg			
c/deg	10.7	7.0	4.8	3.1
22.5	+.36	-.18	+.17	-.06
15.5	+.47	-.29	+.24	-.15
10.7	<u>0.5</u>			
9.5	+.57	+.14	-.10	0
8.1	+.18	+.52	-.21	+.06
7.0		<u>0.5</u>		
6.3	-.12	+.46	+.24	-.10
5.4	-.03	+.34	+.37	-.14
4.8			<u>0.5</u>	
3.8	+.20	-.25	+.52	+.20
3.1				<u>0.5</u>
2.2 *	-.15	+.26	-.52	+.62
1.5 *	-.26	+.42	-.48	+.66
.93 *	-.32	+.57	-.64	+.67
.57 *	-.32	+.58	-.68	+.55
.30 *	-.34	+.55	-.64	+.58
.16 *	-.37	+.54	-.69	+.46

* = Test contrast reduced below 0.5 to mark match.

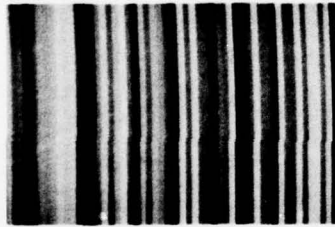
N = 2



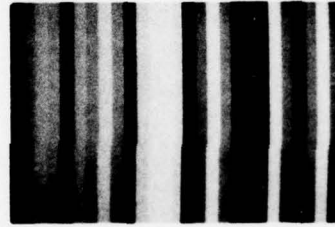
2



4



8



16

Figure 2.9

Texture matches using bars of fixed width (as shown beneath the figures), but variable gray levels. The left half of each picture contains randomly chosen grays. The right half has only three gray levels (.16, .50, .80) chosen with equal probability.

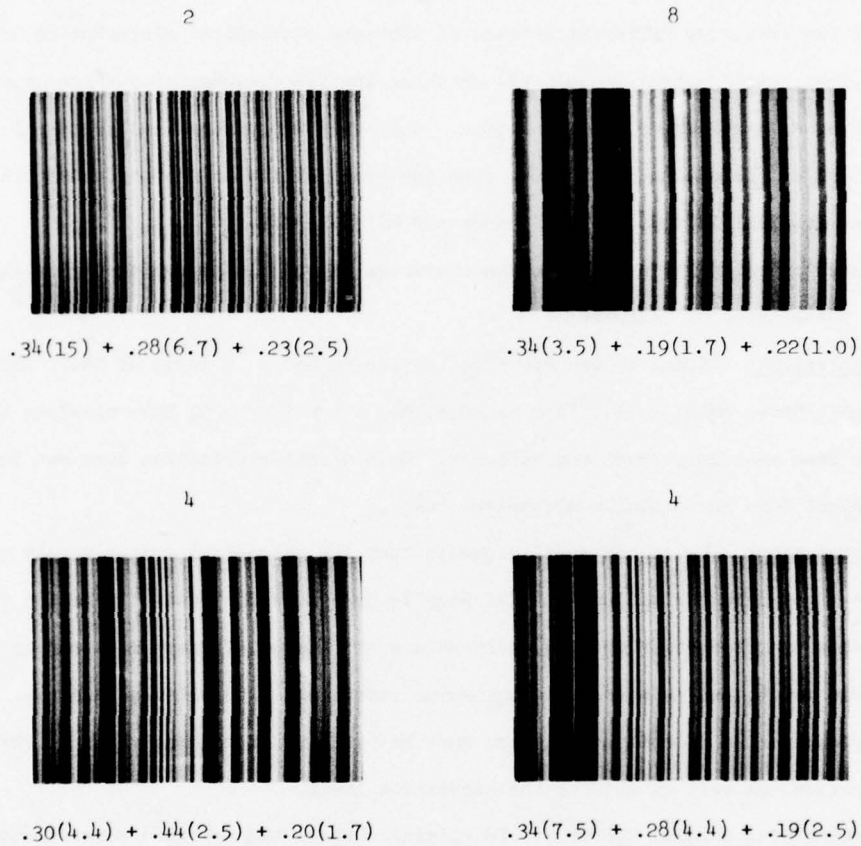


Figure 2.10

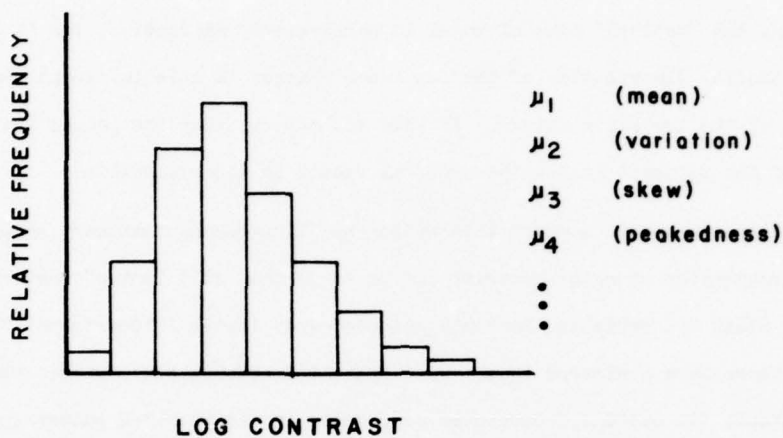
Texture matches using bars of fixed width (as shown above the figures), but variable gray levels. The left half of each picture contains randomly chosen grays. The right half is constructed from three spatial frequencies (given in parentheses) at the contrasts shown.

metamers may also be obtained and that the solutions are not unique. For example, the lower pair illustrates two different matches to the same statistical distribution of random grays. Thus, either three gray levels or three spatial frequencies suffice to match patterns having a large number of gray levels. Only when two dimensional textures (such as checkerboard patterns) are constructed does the exact choice of the three matching gray levels become critical (Riley, 1977; Richards and Riley, 1977).

There are several ways in which we can characterize what is happening during the analysis of these types of textures:

- A) Statistical: First, we can describe texture metamers in terms of their statistical equivalences (Fig. 2.11). For example, all the metamers we have examined have the same mean gray level and variance. Such a characterization does not lend much insight into the probable mechanism.
- B) Compression: The second model suggests that the perceptual system merely compresses the input-output function--such as Werblin describes for retinal function (1970). Such a transformation would clearly reduce the number of grays required in a metameric match, and is a possible mechanism consistent with known physiology. In an implementation of this model, care must be taken to choose the slope of the Gamma function, as well as setting the adaptation level.
- C) Thresholding a Local Operator: To eliminate these two previous constraints, we can modify the Werblin model so that it acts locally and allows only three response states. The operator would examine gray level changes at the boundaries and note the direction of luminance change, providing the change exceeds a threshold. The outputs across an edge would be limited to "black, gray or white." This is a type of "retinex" model, and one version has already been implemented by Marr (Vis. Res., 1974).

"Statistical" Model



"Retinex"

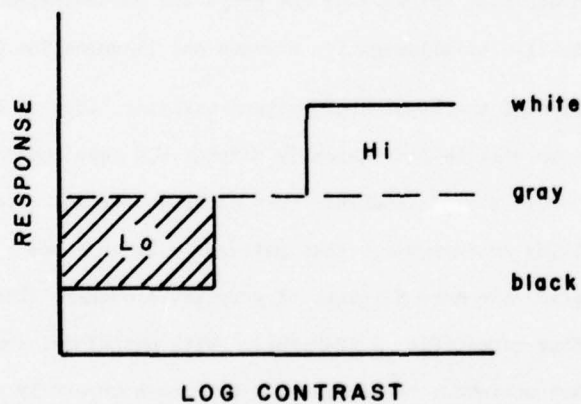


Figure 2.11

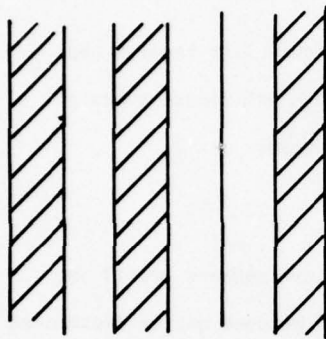
Two models for characterizing the constraints on texture metamers. A statistical model (upper) merely describes the equivalences in terms of the statistical movements. A retinex type model (lower) proposes that the texture equivalences arise from a thresholding operation that segregates darkness from lightness.

To us, the "retinex" type of model is particularly attractive, for it suggests that the spatial distribution of the luminance changes is more important than the actual magnitude of the luminance change. If this is correct, then the actual spatial frequency content of the patterns is not the critical factor in discrimination.

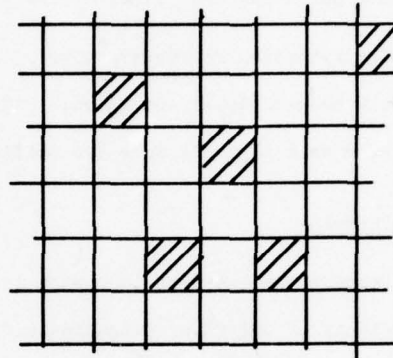
Preliminary results support this notion by illustrating that many triplets of spatial frequencies of equal contrast can be found that will "match" textures having fixed bar width but variable, randomly selected gray levels (Richards and Riley, 1977). Although there is a preferred square-wave spatial frequency that matches the bar width, the additional two spatial frequencies components in the matching pattern, as well as their contrasts, can vary considerably. (See lower pair of matches in Fig. 2.5)

Why are three gray levels or three spatial frequencies sufficient to match "noise" textures of this kind? One tentative answer would be that the object of the system is to recognize textures by their spatial configurations, and not by their precise gray level content. Such a process that throws away the grays and concentrates on the pattern itself would then be insensitive to illumination changes and illumination gradients.

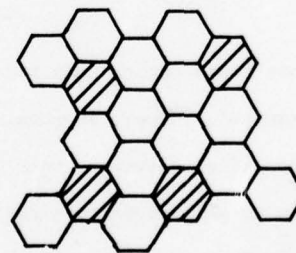
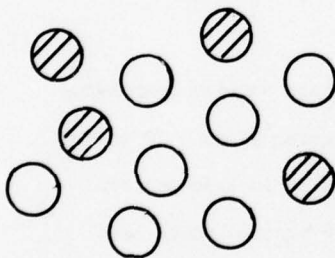
If this is the objective of the high level pattern analyzer, then we might expect that one-dimensional patterns will be intrinsically simpler and less restrictive than two-dimensional patterns, where pattern analysis must occur at several orientations. Thus, it may not be surprising in retrospect that patterns with only one boundary around the elements (Fig. 2.12-left) are more tolerant of gray level changes than patterns having elements with abutting edges (Fig. 2.12-right). With additional abutting edges such as for the checkerboard patterns, the grays must be chosen carefully if only three are to suffice.



3 GRAY LEVELS (many)



3 GRAY LEVELS (few)



?

Figure 2.12

Pictorial summary of effect of texture structure upon the number of gray levels required to match a multi-level pattern of similar structure. The number of abutting edges appears to be the critical parameter.

With more complex arrays, will the minimum number of gray levels required to create texture metamers increase above three? Perhaps so, but as a conjecture let us propose that no more than four grays will be needed, provided that we consider only surfaces and that the structural basis of the pattern is not altered.

The intuition for this number of course comes from the four-color theorem, so recently proved (Appel and Haken, 1977), where only four colors are needed to color countries on a map. If four are sufficient, why bother with more?

vii) Linearity

Perhaps the two most important issues raised by our texture matches are 1) the sufficiency of only four primaries and 2) the assumption that matches may be decomposed into components that can be linearly transformed from one set of primaries to another. These points are difficult to test rigorously, as witnessed by the difficulty in obtaining conclusive answers to these questions even in the rigorous science of colorimetry (Wysecki and Stiles, 1967). The sufficiency of four primaries we know will fail if texture matches are expanded to include all deterministic textures or repetitive patterns.

In some cases information regarding the relative phase of the Fourier components will be required, thereby increasing the dimensions of the matching space. If phase is to be specified, however, then the problem has been enlarged to include pattern recognition as well as texture perception. At present, we have ignored phase (see Atkinson and Campbell, 1974; Hamerly et al, 1977; Sansbury, 1977).

The assumption that the Fourier components of a pattern can be added and subtracted algebraically fails if the contrast of the textures approaches unity, because of

nonlinearities introduced by the saturation of neural activities. At the other extreme, the linearity assumption appears to be approximately valid at low contrasts, but may fail near threshold if threshold setting operations are present (Limb and Rubinstein, 1977; Wilson, 1978). The problem is to determine the extent of additivity failure for various levels of contrast and for a wide range of complex texture displays. Our first step in this direction has been to test the predictions for transforming from one set of primaries to another.

For square-wave primaries, a different set of spatial frequencies was found to be optimal as compared with sinusoidal primaries. Part of the reason for the change in primaries for texture patterns containing sharp lines and edges may be that the texture analysis occurs in a different manner, as suggested earlier. If, in fact, the mechanism for analyzing the distribution of contrast steps is different from the mechanism analyzing the more global spatial content of a pattern, then the two sets of texture matching functions should be partially independent. Specifically, one should not be derivable from the other by a linear transformation of the matching functions.

In the case of transforming the sinusoidal matching functions of Table II into the comparable set of square-wave matching functions of Table IV, the difficulty is obvious. The square-waveforms of the Table IV frequencies must be decomposed into their sinusoidal harmonic contents, each of which must be considered when transforming from the sinusoidal matching functions of Table II. The cumulative errors in such transformations would prohibit a strong test of linearity without exhaustive measurements with small errors.

As a first check on linearity, therefore, texture matches were obtained using the same primaries as the square-wave functions of Table IV, but with sinusoidal waveforms throughout. Thus, we begin by asking the simpler question of whether the fundamental

41

frequency of the square-wave matching set can be predicted from the original sine-wave matching functions of Table II. These new data obtained with sinusoids and without eye movements using the customary 3° wide x 2° high field are given in Table V for two observers. In parentheses next to each measured matching function value is the value predicted by transforming the data of Table II. The agreement is poor, especially considering the fact that the transformation allowed the use of any value of the matching functions within one-third the range found. Even with this very relaxed test, serious linearity failures were found for all test frequencies above 3.1 c/deg as indicated by the asterisks.

More surprising are the results of Table VI. Here the sinusoidal matching functions of Table II have been transformed to approximate the empirical square-wave functions of Table IV, but the difference in waveform ignored. Now there are only four linearity failures! Clearly the transformation from a low frequency set of primaries to a new set of high frequency primaries must be a non-linear transformation.

Thus, the Linearity Property does not hold for texture matches. However, it should be recognized that the new set of primaries used to obtain the matching functions of Tables V and VI represent extreme transformations. Furthermore, the matches of both Table V and VI are very desaturated (they appear like high frequency "white noise"). The sinusoidal functions were generally not as satisfactory as those matches made to construct Table II. (Specifically, the percent of matches judged "very good" or "excellent" was reduced from 90% in Table II to 80% in Table V.)

For small changes in the set of sinusoidal primaries used in Table II, pilot studies show that linearity will hold. This finding implies that the primaries of 10.7, 6.3, 3.2 and 0.93 c/deg lie near the peak sensitivities of the underlying response functions (see section VI).

TABLE V
TEXTURE MATCHES TO VERTICAL SINUSOIDS
(3x2 deg field--no eye movements)

TEST FREQUENCY	PRIMARY FREQUENCY, c/deg			
	10.7	7.0	4.8	3.1
c/deg	Measured (Predicted)	Measured (Predicted)	Measured (Predicted)	Measured (Predicted)
22.5	+16 (.22)	-.03 (-.07)	+.03 (+.03)	-.02 (-.02)
15.5 *	+.35 (.35)	-.13 (-.13)	+.12 (+.01)	-.10 (+.03)
10.7	<u>0.5</u> (<u>0.5</u>)			
9.5 *	+.51 (.51)	+.11 (.11)	-.06 (+.01)	0 (-.02)
8.1 *	+.12 (.23)	+.45 (.42)	-.09 (-.05)	+.01 (-.04)
7.0 *		<u>0.5</u> (<u>0.5</u>)	(+.14)	
6.3 *	-.02 (-.06)	+.54 (.45)	+.09 (.25)	-.08 (-.08)
5.4 *	-.05 (-.06)	+.13 (.23)	+.55 (.33)	-.16 (-.08)
4.8 *		(.18)	<u>0.5</u> (<u>0.5</u>)	
3.8 *	+.12 (+.06)	-.11 (-.11)	+.27 (.27)	.32 (.32)
3.1				<u>0.5</u> (<u>0.5</u>)
2.2	-.10 (-.10)	+.16 (+.16)	-.27 (-.30)	+.56 (.66)
1.5	-.12 (-.12)	+.20 (+.20)	-.29 (-.31)	+.48 (.55)
.93	-.16 (-.25)	+.30 (.50)	-.35 (-.50)	+.44 (.44)
.57	-.16 (-.17)	+.20 (.31)	-.28 (-.28)	+.30 (.30)
.30	-.12 (-.15)	+.21 (.25)	-.32 (-.32)	+.26 (.26)
.16	-.15 (-.11)	+.15 (.15)	-.25 (-.14)	+.18 (.12)

* = Linearity test failure

N = 2

TABLE VI
LINEARITY TEST: SQUARE-WAVE MATCHES (IV)
PREDICTED FROM SINUSOIDS (II)

TEST FREQUENCY	PRIMARY FREQUENCY, c/deg			
c/deg	10.7	7.0	4.8	3.1
22.5	+.34	-.18	-.10	-.04
15.5 *	+.44	-.24	+.02	-.01
10.7	<u>0.5</u>			
9.5 *	+.57	+.14	0	-.02
8.1	+.20	+.50	-.13	0
7.0		<u>0.5</u>		
6.3	-.12	.46	.20	-.10
5.4	-.07	.34	.37	-.08
4.8 *	-.02	+.07	<u>0.5</u>	
3.8 *	+.08	-.25	.52	+.20
3.1		-.05		<u>0.5</u>
2.2	-.48	+.22	-.52	+.76
1.5	-.40	+.42	-.50	+.66
.93	-.50	+.60	-.70	+.65
.57	-.33	+.58	-.54	+.48
.30	-.28	+.55	-.50	+.30
.16 *	-.12	+.23	-.19	+.17

*Significant discrepancy between derived and measured contrasts

viii) Reduced Texture Space

Rather than plotting the actual contrast of a primary required in a match, the contrast ratios between primaries may be specified. In colorimetry, these ratios are designated as chromaticities, and merely correspond to the projection of the tri-stimulus functions upon the (1,1,1) plane. In addition to the simplification introduced by skirting one variable, the advantage of such a projection is that the relations between the primaries are more obvious and appear better correlated with the actual color perception, where ratios rather than magnitudes are more relevant. With these same considerations in mind, Fig. 2.13 illustrates the three dimensional projection onto the (1,1,1) plane of the four dimensional texture space. The x, y and z axes are primaries 5, 2 and 9.5 c/deg, using data obtained from an earlier method (Richards and Polit, 1974). The x, y plane is indicated by the cross-hatched parallelogram. The projection of the frequency locus onto this plane is shown by the solid line, which takes the form of a helix. The dotted line is the *three dimensional* representation, which reaches a minimum in the z direction near 3 c/deg. White noise would lie somewhere near the middle of the loop of the helix.

One of the interesting properties of this texture ratio space is the doubling back of the frequency locus upon itself. In fact, this looping back occurs twice, once near 3 c/deg and again near 0.1 c/deg where the ratios change sign and become infinite. The practical significance of this behavior is not yet obvious to us. One possibility, however, is that the reduplication of ratios even in two dimensions would facilitate the scaling of texture patterns by the visual system as the reference metric was changed from fine to coarse (as during size constancy).

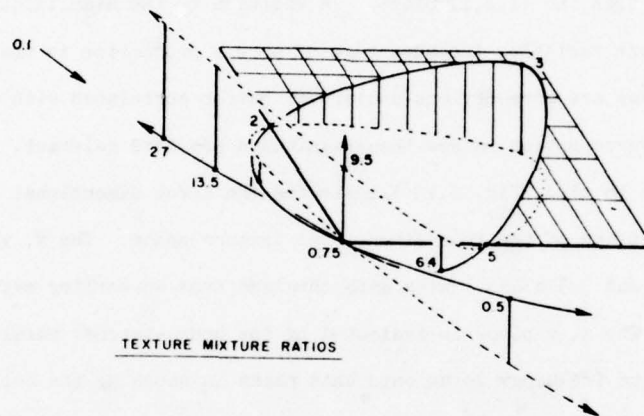


Figure 2.13

Projection of texture matching functions onto three-space, indicating the proportions of each primary needed to match any test frequency. The solid line shows the frequency locus projected onto the x, y plane defined by the 5 and 2 c/deg primaries. The dashed line is the same frequency locus in three dimensions, with the Z-axis defined by the 9.5 c/deg primary. White noise would be located roughly in the middle of the loop of the helix.

III. Two Dimensional Texture Matches

1. Basic Contrast Sensitivity Data

As early as 1966, Kelly reported spatio-temporal sensitivities for Bessel functions, and more recently, the sensitivity to stationary orthogonal sinusoidal profiles have been examined (Carlson et al, 1977; Burton, 1976) as well as filtered noise patterns (Mostafavi and Sakrison, 1976; Mitchell, 1976; Koenderink and van Doorn, 1974). An important type of two dimensional stimulus not yet studied, however, is the simple Fourier generalization of the temporally modulated, sinusoidal luminance profiles originally used by Robson in 1966.

Following Robson, we first generated two dimensional stationary luminance profiles L_s , that were represented by

$$L_s = L_o (1 + m_x \cos 2\pi V_x x) \cdot (1 + m_y \cos 2\pi V_y y) \quad (1)$$

where m_x, m_y are the contrasts in the orthogonal x and y directions, and V_x, V_y the spatial frequencies. Temporal modulation of this pattern was then introduced by rearranging the expansion of the above equation into two components, one consisting of the sums and the other of the products of the cosine terms:

$$L = L_o \{ 1 + G_s(2\pi ft) \cdot (m_x \cos 2\pi V_x x + m_y \cos 2\pi V_y y) + G_p(2\pi ft) \cdot (m_x m_y \cos 2\pi V_x x \cdot \cos 2\pi V_y y) \} \quad (2)$$

The temporal modulation function, G, was a maximally modulated sine, triangular, or a square-wave function, as indicated later.

Two basic types of patterns were of interest, each symmetrical in x and y with $V_x = V_y = V$ and with $m_x = m_y = m$. The first displayed only the sums of the individual spatial cosine functions in x and y of the equation (2) and set the product term equal to zero (i.e. $G_p = 0$). The second and complementary type of pattern presented only the products in x and y and set the sums of the two cosine functions to zero (i.e. $G_s = 0$).

These two terms are orthogonal functions and together represent the important components of the basic Fourier trigonometric system in two variables. It is the behavior of this product term, however, that is of particular interest, for the human spatio-temporal sensitivity to such a profile has not yet been examined. Yet this pattern is the basis for constructing the checkerboard patterns so commonly used in evoked-potential studies.

Both types of symmetrical patterns were set up on a 440 x 440 Video display with the modulation amplitudes, spatial and temporal frequencies controlled with a PDP 11/10 computer plus some associated hardware. The refresh cycle of the system was 33 msec because of raster interlacing, and thus temporal modulation was square-wave for 15 Hz, triangular wave for 3.8 Hz. For each point, at least three separate measurements were made binocularly by the author at a distance of 2 m. Reproducibility was within 10%. The grating patterns subtended $5^\circ \times 5^\circ$ in a dimly lit room with the mean luminance of the display at 20 cd/m^2 . A fast P4 phosphor was used to produce a broadband white stimulus.

Figure 3.1 shows that the general form of the sensitivity functions for both the product and sum terms are similar to those found by Robson in 1966. Both of these functions, like Robson's, exhibit two features. First, the form of the fall-off in sensitivity at high spatial frequencies is independent of temporal frequency, and is similar for both the sums and products provided that the product sensitivities are plotted as $1/m^2$ to correct for the additional peak-to-trough signal attenuation introduced by multiplying the x and y luminance distributions. Second, a marked fall-off in sensitivity at low spatial (or temporal) frequencies occurs only when the temporal (or spatial) frequency is low (circles).

Several additional features may also be noted in these functions. First, the function describing the sensitivity to the stationary sum of the two orthogonal gratings (circles) is the same as the sensitivity to either a vertical or horizontal grating presented alone

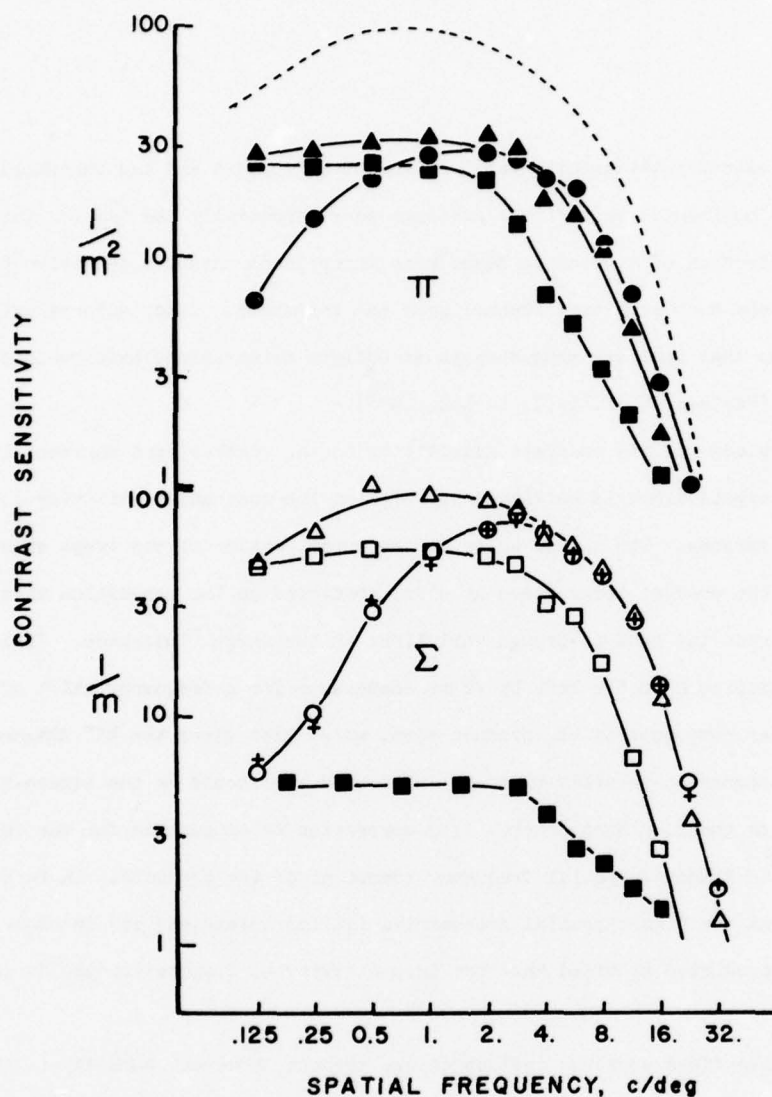


Figure 3.1

Spatial contrast sensitivity functions (reciprocal of threshold modulation, m) for different temporal frequencies. Upper set of curves are threshold functions for the products of sinusoidal luminance gratings in x and y , and modulation squared is plotted. The lower set of functions are for the sums in x and y . The plusses are thresholds for one-dimensional sinusoids (average of vertical and horizontal). Circles: 0 Hz; triangles: 4 Hz; squares: 16 Hz. Filled symbols represent product thresholds, open symbols represent thresholds for the sums.

plusses). (The author's astigmatism was less than 0.5 diopters and the threshold measurements for either horizontal or vertical gratings were essentially the same.) This result suggests that detection of the sum is based upon independent channels sensitive to orientation and that the most sensitive channel sets the threshold. Other authors using different wave forms that included measurements at oblique orientations have reached this same conclusion (Carlson et al, 1977; Burton, 1976).

Second, knowledge of the contrast sensitivity to the vertical and horizontal sinusoidal grating targets alone is sufficient to predict the contrast sensitivity to the product of such targets. The dotted curve at the upper portion of the graph shows the sensitivity for the product terms (seen at 4 Hz) predicted on the assumption that detection is based simply upon the peak-to-trough variations in the target luminance. (This dotted curve has been displaced to the left by $\sqrt{2}$ to compensate for a frequency shift of the fundamental linear component of the product term, which lies along the 45° diagonals.) In this case the contrast sensitivity to the x, y products should be the square-root of the sensitivity to the x, y sums after a $1/\sqrt{2}$ correction to compensate for our reduced sensitivity to the diagonal spatial frequency component of the products. In fact, the measured values at the higher spatial frequencies (filled triangles) are in close agreement with those predicted provided that the insensitivity to diagonal sinusoids is taken into account.

When receptive field surround influences are present, however, such as at 0 Hz flicker and low spatial frequencies, then the square-root of the product thresholds (filled circles) begin to approximate the sum thresholds (open circles). This equivalence between the sum and product thresholds implies that the threshold-setting mechanism for the surround is not orientation sensitive. At low spatial and temporal frequencies, therefore, the low frequency fall-off may be a property of the simple, circular center-surround receptive fields typical of retinal ganglion cells (Kuffler, 1953).

2. Texture Matches

At threshold, the contrast sensitivity for detecting sinusoidal gratings is independent for horizontally and vertically oriented patterns (Carlson et al, 1977; Burton, 1976; Richards, previous section). Thus it was not surprising to find that all previous texture matches made to vertically oriented patterns were valid when an identical horizontal pattern was added to create a two dimensional "plaid" texture (see Fig. 3.2a). In some cases the quality of the texture match actually improved by adding the orthogonal components. Only seldom was the texture match impaired.

A second type of two dimensional pattern can be created by multiplying, rather than adding, the waveforms in x and y. (See Fig. 3.2b). Such a pattern contains the products of the spatial frequencies of the x and y components, and leads to pattern components at many different orientations (Kelly and Magnuski, 1975; Kelly, 1977). In the reduced case where only one spatial frequency is presented in x and y, the product is a 45° oriented sum of a frequency 1.4x the original. It is generally not possible to match even this simple pattern by primary components whose orientation is confined to horizontal and vertical (0 and 90°). Furthermore, as previously mentioned, the spatial frequency primaries for textures oriented at 45° must be considered for texture matches to the most general patterns containing components at all orientations. Whether or not four orientations are sufficient for all possible textures has not yet been determined, although preliminary data obtained by M. Riley (1977) shows that four orientations are sufficient for line elements of equal length but random orientations. (See also Fig. 6.2).

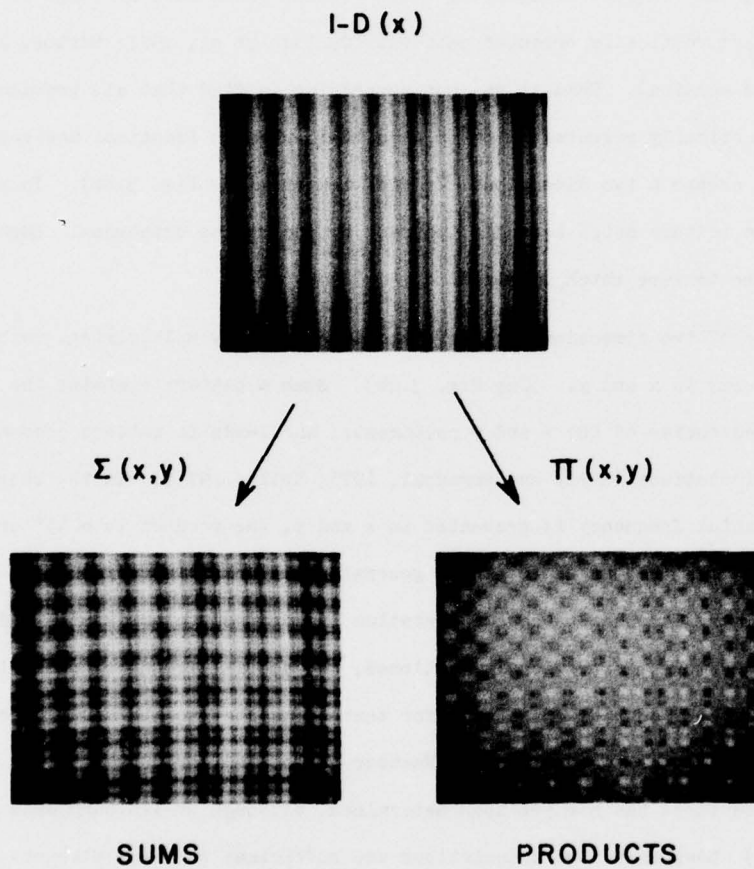


Figure 3.2

Examples of the sums and products of a complex texture pair (top).

IV. Random-Dot Textures (with S. Purks)

N-GRAM PATTERNS

In the early sixties Julesz (1962) created computer-generated patterns with controlled high-order statistical properties. Although these patterns for the most part appear as random collections of dots, the patterns assume different textures as the statistical properties of the dots are altered. However, the texture of these patterns is unfamiliar. Therefore, Julesz argues, when subjects are asked to discriminate such stimuli, they should be forced to use their more primitive visual mechanisms.

If Julesz's position is correct, then the use of random-dot patterns with controlled statistical properties is one method of revealing some basic organizing principles of primitive information processing. Julesz's results suggested that the discrimination of random-dot texture patterns was based primarily on the analysis of clusters or lines formed by proximate points of uniform brightness.

Textures with different size clusters or runs of points of equal brightness may be generated by controlling the probability transitions of adjacent pairs of elements. A key question, therefore, is whether the statistics of the pattern is the index of discriminability, or whether other measures, such as the spatial frequency content, provide more appropriate indices. Julesz's earlier work suggested that n -gram patterns with $N > 3$ were less discriminable than patterns based on one- or two-gram statistics. However, Fig. 4.1a is an example of two discriminable patterns that differ only in their statistics for four adjacent points (4-gram). This discrimination of patterns despite identical 1-gram and 2-gram statistics raises again the question of the relation between statistical complexity and the nature of visual discrimination. More explicitly, is there a direct relation between statistical dependencies and the process of visual discrimination? To answer

this question, a new method was devised for generating patterns with controlled n-gram statistics. Then the method was applied to isolate the effects of manipulating the statistics of sequences of various lengths (or spans), while leaving invariant the statistics of shorter spans.

METHOD OF GENERATING TEXTURES

Binary sequences with transitions dependent on the $n-1$ previous points were generated on a PDP-8 or PDP-12 computer, and then were translated and read out on a Calcomp plotter to create the white = 0 and black = 1 squares which made up the visual patterns. Each sequence was 2048 elements in length and was sliced and plotted as a rectangular half of a 64×64 array. Generally, the slices were 64 elements long laid parallel to the division between the two halves, except for control figures where 32-element slices were laid perpendicular to the division. A second sequence with different n-gram statistics filled the second half of the array. To control for orientation preferences, all patterns were tested for discriminability with the divider oriented both vertically and horizontally; this was accomplished simply by rotating the pattern.

To test for the discriminability of patterns that differ only in their statistics for spans greater than n , two separate sequences must be generated that differ in their n-gram statistics, but which leave all shorter span statistics identical. To accomplish this task, we have proven elsewhere that the generation of probabilistic sequences of length n may be defined completely by a set of transition probabilities for each of $2^{(n-1)}$ possible previous subsequences of length $n-1$ (Purks and Richards, 1977).

The proof shows that for a variable V_j which takes on the binary values 0 or 1, the apriori probability of a sequence of length n defined as $P(V_1, V_2 \dots V_n)$ may be divided into two probability functions, G_1 and G_0 , for generating either a "1" or a "0" following the shorter sequence of length $n-1$. The relation between G_1 and G_0 and the sequence probability functions is as follows:

$$\begin{aligned}
 &P(V_2, V_3, \dots, V_{n-1}, 1) \\
 &= G_1(1, V_2, V_3, \dots, V_{n-1}) * P(1, V_2, V_3, \dots, V_{n-1}) \\
 &\quad + G_1(0, V_2, V_3, \dots, V_{n-1}) * P(0, V_2, V_3, \dots, V_{n-1}) \quad (1)
 \end{aligned}$$

Note that fixed (n-1) gram constraints define values of the form $P(V_1, V_2, V_3, \dots, V_{n-1})$ in the above equation. Thus the fixed (n-1) gram statistics will determine a particular dependence between pairs of generation parameters of the form $G_1(1, V_2, V_3, \dots, V_{n-1})$ and $G_1(0, V_2, V_3, \dots, V_{n-1})$.

ILLUSTRATION OF GENERATION METHOD

To illustrate the convenience of the above method for controlling n-gram statistics, consider Fig. 4.1b. Each half of this figure has identical 1-gram and 2-gram statistics with $P(V_1, V_2) = 0.25$ for all permissible values of V_1, V_2 . The top and bottom halves of the pattern are clearly discernable, however, because there are substantial differences in their statistics for spans of length 3 or more. To hold the 1- and 2-gram distribution constant in both half-fields, the following generation parameters were used:

Bottom Field	Top Field
$G_1(00) = 0.95$	$G_1(00) = 0.05$
$G_1(01) = 0.95$	$G_1(01) = 0.05$
$G_1(10) = 0.05$	$G_1(10) = 0.95$
$G_1(11) = 0.05$	$G_1(11) = 0.95$

Note that in each half-field, $G_1(00) + G_1(10) = 1$ and $G_1(01) + G_1(11) = 1$, or more generally, that

$$G_1(0, V_2) + G_1(1, V_2) = 1.$$

Thus, Eq. (1) is satisfied for uniform 2-gram statistics.

TEST PROCEDURES

Following pilot studies, nine patterns were chosen for discrimination testing on six subjects (four naive plus the two authors). Each pattern was 5 in. square and contained two subfields of equal size that divided the square either right-left or top-bottom. Both orientations were used for each pattern. The task of the subject was to report first the orientation of the division between the fields, i.e., top-bottom or left-right. Then he assigned a number to indicate the magnitude of the discriminable difference between the two subfields. The experimenter recorded this magnitude estimate as a negative value if the subject's perceived division disagreed with the actual objective division.

To assist each subject in assigning numbers indicating the discriminability of the texture pairs, three reference patterns were in constant unobstructed view at 1 m distance, subtending 8° visual angle. The first was completely random throughout and was assigned a scale value of 0. The second, which was identical to the tested pattern displayed in Fig. 4.1c, was assigned a value of 1. A third picture (Fig. 4.1b) with still greater differences between the 3-gram statistics of each half was used to define the scale value 2 (see Purks and Richards, 1977 for the generation probabilities). Subjects were encouraged to use fractional scale values, and of course, could also use numbers larger than 2 if appropriate.

In addition to presenting patterns with different statistical properties at the 1 m viewing distance, several experimental manipulations were introduced. These included (i) changing the viewing distance, (ii) blur, (iii) changing orientation to 45° , and (iv) tilting the pattern out of the frontal plane. For all of these manipulations, the average luminance of the patterns remained in the photopic range near 250 cd/m^2 .

Figure 4.1



- A. Texture pattern where upper and lower halves of figure differ only in their 4-gram statistics.



- B. Texture pattern with identical 1- and 2-gram statistics in the top and bottom half. This is a scale reference figure having the value 2.



- C. Reference pattern having a value of 1. The 1- and 2-gram statistics are constant, whereas run length only has been controlled by a suitable variation of the 3-gram statistics. The top field favors runs of length two, while the bottom field favors runs of one or more than two.



- D. Controlled 3-gram statistics. The spatial frequency content of both half-fields is the same, but the two fields have a difference in their 3-gram statistics which is comparable to that in Fig. 4.1c. The top and bottom halves of the figure differ only in their phase relationships.

RESULTS

Control of run-length

From Julesz's earlier studies, it was clear that if the average gray level of two patterns was identical, then the distribution of black and white run lengths appeared to be the most important variable in aiding texture discrimination. Run length may be controlled by varying the parameters controlling 3-gram generation while holding the 1- and 2-gram generation parameters constant. Figure 4.1b is such an example. The top field of this figure favors runs (extent of uniform lightness) of length two, while the bottom field favors runs of either one or greater than two. Note that the overall impression is one of a difference in "coarseness" of the two subfields. Such a difference may also be correlated with the distributions of spatial frequencies contained in the two fields, and hence a possible basis for the discrimination is one based on a spatial-frequency analysis.

Because there are two independent generation parameters for 3-grams with fixed 2-gram statistics, it is possible to control both black and white run length independently. In Fig. 4.1 c, run lengths of both black and white are kept short in one half-field and are prolonged in the other. The two half-fields thus differ in their spatial-frequency content and are clearly discriminable. The magnitude of this difference on our scale is 1.

If now a new pattern is created with identical 1- and 2-gram statistics but alternate 3-gram statistics that do not generate a difference in the spatial-frequency content between the two half-fields, then discriminability is essentially lost. Figure 4.1 d is such an example. Its 3-gram statistics differ from Fig. 4.1 c by an inversion of probabilities when bit 2 is zero. This is a result of changing the generation probabilities in order that the top field will favor long white run lengths plus short black run lengths, while the bottom field will favor short white plus long black run lengths. Thus the two fields differ only in their phase relations; and complementing the black-white

elements in the top field will yield a pattern statistically identical to that of the bottom field. Although Figs. 4.1c and 4.1d each have sub-fields with equivalent differences in their 3-gram statistics their discriminability is quite different. This result suggests that the n-gram statistics are not the primary variable of interest of the visual analyzing mechanism.

Figure 4.1a reinforces the independence of the basis for visual discrimination and the span of the n-gram. In this figure, an obvious separation between half-fields of magnitude 1, 2 is achieved by varying only 4-gram statistics. Once again, like the previous figures, this discriminable difference is present regardless of the orientation of the pattern, whether the border between the half-fields is vertical, horizontal, or inclined at 45° .

In all of the previous figures, differences in the average length of black or white runs promoted discrimination between the half-fields. Such runs are controlled largely by the 2-gram statistics. With controlled 4-gram statistics it is possible to keep the distribution of run lengths constant because there are four independent generation parameters, only two of which affect run length distribution. Figure 4.2a is such an example where the run length distributions are equivalent to "chance." However, the right half-field favors "symmetric" alternations such as 01010 or 110011, while the left half-field favors asymmetric alternations such as 11011 or 00100. Thus the half-fields do not differ in their run length distributions but do differ in their spatial-frequency contents. For a vertically or horizontally oriented border between the half-fields, the difference is discriminable at a scale value of 0.5. (Incidentally, this discrimination improves markedly if the field is reduced by viewing at 5 m, whereas orientation at 45° degrades the differences.) Thus the distribution of run lengths is not a necessary basis for

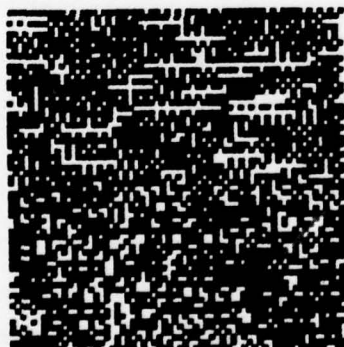
Figure 4.2



- A. Texture patterns with the distribution of run lengths held constant by controlled 4-gram statistics. Discrimination is possible, having an average scale value of 0.5. The right and left of the figure differ in "symmetry" of the alternation of black-white transitions.



- B. Controlled 4-gram statistics as in Fig. A except that the directionality of black-white transitions is manipulated. No discrimination of top and bottom subfields.



- C. Texture patterns with the black squares twice as frequent as the white, thereby changing the gray level. The run lengths in each half of the pattern are controlled by 3-gram statistics, as in Fig. 4.1c. Discrimination of the top and bottom half-fields is poor, with a value 0.1.



- D. Identical to Fig. C except that the white points have been complemented to black and vice versa. Discrimination is now increased to 1.1.

discrimination, but the arrangement of runs affecting spatial-frequency content can be critical.

As still another example of the difficulty of predicting discrimination from n-gram statistics, consider Fig. 4.2b . Here again run length is chance as in the previous Fig. 4.2a, and only the alternation sequence is manipulated as before. But in this case the two half-fields differ only in the direction of black-white transitions. For example, the top field favors sequences such as 010011, whereas the bottom half-field favors 110010. Here the two fields have similar spatial frequency contents but differ in their phase preferences. No observer could reliably make this discrimination (scale value equalled 0), suggesting again that local phase information is lost during texture analysis.

Global variations in spatial -frequency content

If subjects are asked to analyze the basis for their discriminations, all invariably indicate that the length of black or white runs (or clusters) is the most important clue. Such runs determine the major spatial-frequency components of the pattern, aside from the dot size itself. Clearly, if a given pattern is viewed from a greater distance, then its spatial-frequency content will increase in proportion, whereas its statistical properties will remain invariant (at least as long as the dots may be still resolved). If it is the statistical properties of the display that are important in controlling texture discrimination, then any such change in viewing distance will not change discrimination. In order to make this comparison between discrimination of patterns with fixed statistics but variable spatial-frequency content, several experimental manipulations were introduced.

1. Viewing distance. A change in the viewing distance from 1 to 5m generally altered the magnitude of this discrimination. The largest change occurred for Figs. 4.1c and 4.2a, whose scale values, respectively, increased from 1 to 1.3 and 0.5 to 1.1.

2. Blur. All patterns were also viewed through spherical and cylindrical lenses of ± 6 diopter power in both orientations. Spherical blur generally severely reduced discrimination, with the sole exception of Figs. 4.1c and 4.2a. Cylindrical blur perpendicular to the direction of the runs generally caused a degradation, reducing the magnitude of the difference between half-fields to 80% of its original value without lenses. On the other hand, if blur was introduced parallel to the runs, then discrimination generally increased, often by as much as 50%. The difference due to the direction of cylindrical blur is such that the latter axis eliminated the high-frequency content introduced by the dots in a string, thus "bringing out" the lower-frequency runs by streaking, whereas perpendicular blurring only eliminated the high-frequency spectra created between dots lying above and below each other. Emphasizing the low-frequency content of the strings of a pattern thus aids the discrimination of random-dot textures.

3. Tilt. A further manipulation of the spatial-frequency content of the patterns was obtained by tilting the patterns out of the frontal plane by about 80% (i.e. as near the sagittal plane as possible without loss of dot resolution). With the viewing distance now reduced to about 50 cm, the range of spatial frequencies thus varied as a gradient of about 25%. Figures 4.1a, 4.2a and 4.2d were most affected by this manipulation, with discrimination reduced almost to zero when the tilt was perpendicular to the border between the half-fields, thus creating a gradient in the direction of the runs. No pattern was better discriminated when seen as tilted in any orientation. This result suggests that texture discrimination is easier if the spatial-frequency content of texture pairs is constant, at least in the direction of the pairwise analysis.

Nth ORDER PATTERNS

The above results and Figures 4.1, 4.2 all are based on N-gram statistics. The results suggest that analyzing textures in terms of the span of statistical dependencies is not a fruitful approach to understanding human texture perception. This conclusion has implications for patterns having broader classes of statistical dependencies.

Fig. 4.3 is a demonstration that the nature of the statistics (n-gram versus n^{th} order) is not the critical factor for understanding human texture perception. This pattern has identical 1st and 2nd order statistics in each half, but different third order statistics. A dozen subjects easily saw a left-right difference by noting the longer runs in the right half. Of interest is that most observers see only the white or only the black runs at any given instant, with the runs of opposite contrast being merged with the middle gray to form the background. This "figure-ground" effect suggests that features with positive and negative contrast are processed in parallel at some level. (See also Figs. 4.2c and 4.2d.)

To generate Fig. 4.3, we used a procedure similar to Julesz (1962) with the third-order probability distribution given by transition probabilities $P(k/ij)$ as follows:

$$P(k/ij) = P[2k-i-j = S(\text{mod } 3)] = P(S)$$

where i , j , and k are successive samples along a horizontal line. For the left half-field,

$$P(S = 0) = 1/16$$

$$P(S = 1) = 5/16$$

$$P(S = 2) = 10/16$$

while for the right half-field

$$P(S = 0) = 10/16$$

$$P(S = 1) = 5/16$$

$$P(S = 2) = 1/16$$

Here 0, 1 and 2 correspond to the three gray levels that had a reflectance of 0, .33 and .90.

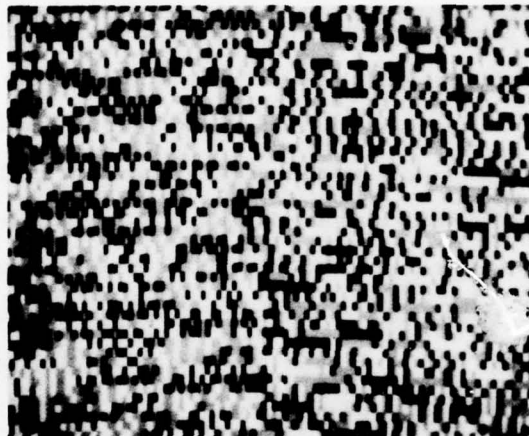


Figure 4.3

A 3rd order texture pattern with identical 1st and 2nd order statistics in the left and right halves.

In the four gray level pattern originally used by Julesz (1962), the second order statistics in each half were uniformly distributed (i.e., flat). In our three gray level pattern the second order statistics are not uniformly distributed but are identical in the two halves. Specifically, the second order statistics for moments corresponding to gray level pairs spanning 3, 6, 9, ... positions are biased in favor of repetitions. Thus the measured frequencies of repeating a gray level three positions further along in the sequence is 0.167 ± 0.006 , whereas the measured frequency of occurrence of each of the two non-repeating gray levels is 0.084 ± 0.005 . (For moments corresponding to spans not a factor of 3, the distributions are flat with a frequency of 0.111 ± 0.006 .) Yet, it is the differences in the third order statistics that permit the discrimination of the two half-fields of Fig. 4.3. These third order statistics are controlled to produce long runs in the right half-field and short runs in the left.

More discriminable patterns can be obtained by using more extreme statistics together with suitable choices of gray levels. Clearly, the selection of the gray levels can be a crucial factor in highlighting the runs, especially if many gray levels are used. For example, clustering all grays but one at one extreme of contrast can help highlight the longer runs in one half of the pattern. Obviously such manipulations are independent of the order of the statistic, and yet are critical factors for human discrimination.

Discussion

Several of the experimental results suggest to us that analyzing textures in terms of the span of their statistical dependencies is not the most profitable approach to understanding human texture perception. Although, as Julesz (1962, 1971) has pointed out earlier, there is a relation between 1- and 2-gram distributions and discrimination, this relation strongly relates to visual perception only for the shortest span lengths that control gray level. These short span dependencies are also the least affected by viewing distance, for

the achromatic variable is intensive, and not extensive as in the case of spatial-frequency content, which is controlled by longer span statistics. Clearly, if the visual system analyzes patterns in a manner dependent upon angular subtense, then the statistics of a pattern become an almost irrelevant variable. That the spatial-frequency content is an important variable in pattern discrimination is already obvious from many previous studies (Robson, 1966; Sachs, Nachmias and Robson, 1971; Graham and Nachmias, 1970; Stromeyer and Julesz, 1972; and Stecher et al, 1973). The several experimental manipulations such as altered viewing distance, and introducing blur, merely are demonstrations of the obvious.

Not so obvious, however, are the effects of the gray level upon discrimination. Why should the complementary patterns of Figs. 4.2c and d be so different in discriminability? Apparently the analysis of visual texture is influenced by the gray level or separately extracts white and black strings, preferring the black against white. Fig. 4.3 also reveals this figure-ground effect. Such a dissociation dependent upon contrast is not without precedent, having been observed first at the single-cell level by Kuffler, 1953, and more recently psychophysically by Spillmann and Levine, 1971, and by DeValois, 1977. Such contrast-dependent differences would require a hierarchical processing by any mechanism that analyzed patterns in terms of their statistical properties, for the gray level set by the short-span dependencies is shown to influence the higher-order analysis.

Several of our experimental patterns suggest that the spatial-frequency variable is important in texture perception. For example, the discrimination of different 4-gram statistics is subjectively based upon the visibility of runs, and this difference is enhanced by properly oriented blurring to cause streaking and to eliminate the high-frequency content introduced by the individual dots. Streaking, of course, produces a stimulus that is more favorable for triggering bar detectors. Thus texture patterns built from line elements such as those introduced by Pickett, 1964, 1968, may provide more insight into

the mechanisms of texture perception than do random dot patterns. Although such patterns have familiarity cues that random dots do not have, their presence will serve to trigger the known feature detectors more effectively, thus aiding in the analysis of their role. Clearly, these detectors play an important function in texture perception, but this function will be overlooked if they are in effect stimulated only by noise.

V. Site of Texture Matching

i) Binocular Matching

For the psychophysicist, an advantage of two eyes over one is that monocular and dichoptic stimulation may be compared in order to determine whether the visual system considers both inputs equivalent. If both types of stimulation are equivalent then the constraints upon the analysis must be cortical, beyond the site of binocular interaction (Julesz, 1971). Using this technique, one can then determine whether the filters responsible for texture equivalences are peripheral or central. Thus in principle, a texture match can be attempted such that at least two of the components in one patch will be presented to separate eyes. If the texture matching functions now are the same as those obtained when all components are presented to only one eye, then the physiologic basis for texture equivalences must be cortical.

Unfortunately, the dichoptic presentation of two vertical spatial frequencies can lead to rivalry or binocular tilt, depending upon the relative spatial frequencies of the two patterns (Blakemore, 1970; Maffei and Fiorentini, 1971; Fiorentini et al, 1976; Richards and Foley, 1978). When tilt occurs, textures with only horizontal components must be used. Nevertheless, pilot studies showed that even with horizontal gratings, rivalry still persists for most combinations of test and primary spatial frequencies whenever the test frequency is below 3 c/deg. The texture clearly does not have the same appearance as the same pattern viewed monocularly without splitting its components for dichoptic presentation. On the other hand, if the test frequencies are higher than 8 c/deg for sinusoids or 3 c/deg for square-waves, then good dichoptic matches are possible. With good fixation and without eye movements, then the contrasts of the primaries are close to those of Table II for sinusoidal gratings, or to Table IV for square-wave gratings. Thus, texture matches to high spatial frequencies and "edges" probably involve cortical mechanisms.

This failure to obtain binocular texture matches at low spatial frequencies does not require a subcortical locus for this component of the texture analyzing mechanism, however. Although such a locus might be favored by the inability to obtain suitable texture matches, it is possible that the dichoptic phase relations between the pattern components are analyzed in a manner that confuse the texture analysis (i.e. as might be caused by horizontal or vertical disparity interactions).

ii) Adaptation

A second test for the locus of sinusoidal texture analysis makes use of adaptation to spatial frequency (Blakemore and Campbell, 1969; Gilinsky, 1968; Pantle and Sekuler, 1968). These authors showed that following adaptation to a pattern of one spatial frequency, contrast thresholds would be elevated at the same and neighboring frequencies. This adaptation technique yielded a narrow band adaptation effect that was taken by the authors as indicating that the visual system decomposed the stimulus into its Fourier components. Although not always explicitly stated, many investigators have assumed that these narrow band channels revealed by the adaptation technique were comparable to bar or line or other feature detectors seen by neurophysiologists in the cortex of lower mammals. The cortical locus of the adaptation is supported by a 70% interocular transfer of the effect in normal stereo observers (Mitchell and Wade, 1975).

At issue is whether these narrow-band "channels" revealed by the adaptation technique are located before or after the texture matching mechanisms. To answer this question, a texture match is set up such that the textures will appear equivalent but the components in each pattern will be different. If the visual system analyzes these patterns in terms of each of its narrow-band cortical components, then adaptation to one of the components should upset the texture equivalences. Thus, the experimental procedure is to adapt a given spatial frequency used to comprise one of the texture patterns. Following adaptation, the

texture match is reset by the subject. If the texture match has been changed, principally by the addition of an increasing amount of the adapting frequency to one of the primaries, then it is clear that the narrow band channel preceded the texture analyzing mechanism. On the other hand, if the amounts of each component remain the same following adaptation to only one of them, then the Blakemore and Campbell adaptation effect must follow the spatial filters that provided the basis for texture matching analysis. Again, the logic for this argument is quite analogous to that used for color vision: whenever a color match is undisturbed by a manipulation such as chromatic adaptation, then it is clear that that manipulation must have affected the visual system after the filtering provided by the color receptors.

Figure 5.1 shows the results of 2 c/sec counterphase adaptation to each of the four sinusoidal primaries used to construct Table II. (The initial adaptation period was 1.5 min followed by cycles of 15 sec adaptation and 5 sec pattern viewing or threshold setting. The adaptation field was $7^\circ \times 6^\circ$ and covered completely the two $3 \times 2^\circ$ test fields.) If the primary frequency is presented alone, the rise in threshold following adaptation to the same frequency is approximately two-fold (open circles), except for the lowest spatial frequency of 1.3 c/deg. These values are consistent but less than those reported by Blakemore and Campbell (1969) as indicated by the filled circles.

The crosses in Figure 5.1 show the effect of adaptation on a spatial frequency when that frequency is part of a texture match. When 1.3 c/deg is the adapted component (Table II test frequency of 4.5 c/deg), there is no adaptation effect, as expected. For the 3.2 and 6.3 c/deg primaries, there is also no adaptation effect when these gratings are presented as a component of a texture match (1.8, 2.2 and 8.1 c/deg test frequencies in Table II). This unexpected result suggests that texture matching precedes or is independent of the narrow-band "cortical" channels revealed by adaptation.

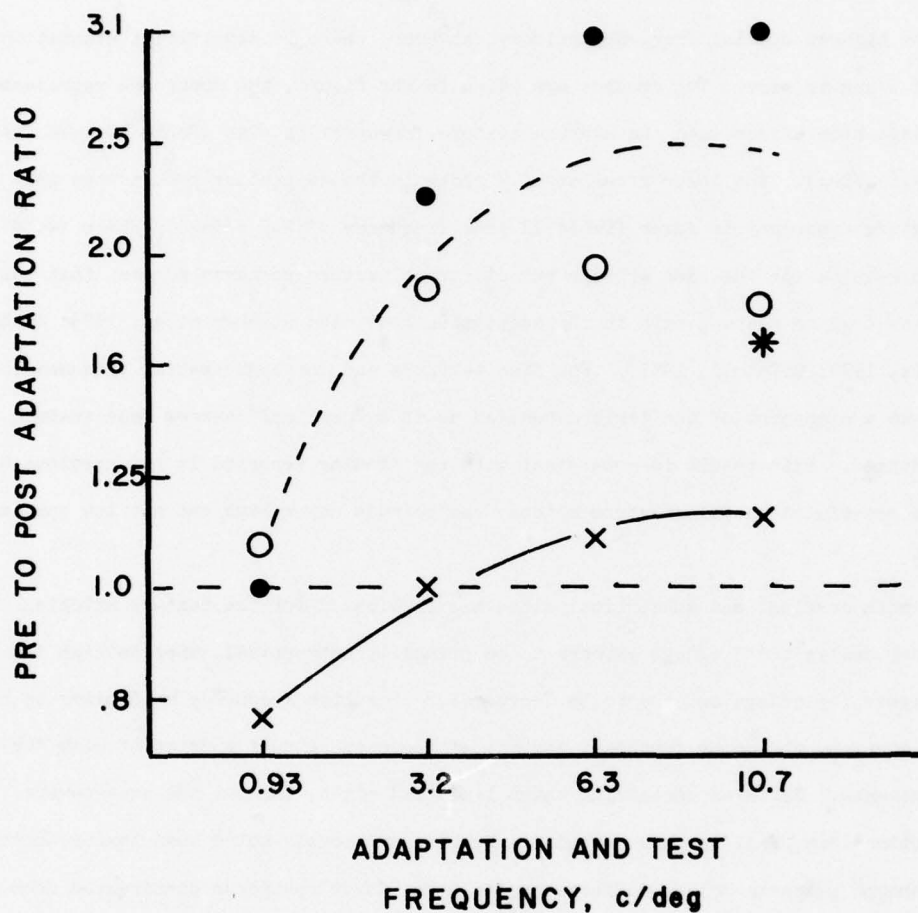


Figure 5.1

Adaptation test for locus of texture matching constraints. Filled circles, Blakemore's adaptation result; open circles, current replication; crosses, test frequency appears as a component in a texture match; asterisk, same as crosses, but a higher test frequency of 4.5 c/deg.

For the highest spatial frequency primary, however, there is significant adaptation to the 10.7 c/deg primary. Two crosses are given in the figure, the upper one representing the mean adaptation effect when the average texture frequency is high (Table II test frequency of 4.5 c/deg). The lower cross at 10.7 c/deg is the adaptation result when the average texture frequency is lower (Table II test frequency of 2.2 c/deg). These two different mean results for the same primary but different texture patterns suggest that the nature of the texture plays a role in the adaptation (see also Stecher et al, 1973; Graham and Rogowitz, 1976; DeValois, 1977). For fine textures and for high spatial frequencies, adaptation to a component of the texture behaves as if a "cortical" narrow band channel had been adapted. This result is consistent with the finding reported in the previous section that successful dichoptic texture matches can be made using high but not low spatial frequencies.

Thus, both cortical and subcortical sites are implicated for the texture matching. Low frequency analysis (<3 c/deg) appears to be primarily subcortical, whereas high frequency analysis (>5 c/deg) appears to be "cortical." For high frequency mechanisms to come into play, however, the major frequency content of the pattern must consist of high frequency components. Patterns containing sharp lines and edges, such as the square-wave grating combinations, fall in this category. We have previously noted that texture matches to these "sharp" patterns follow different rules from "fuzzy" patterns constructed from a few sinusoids. Specifically, the matches to sinusoidal textures emphasize the global spatial frequency content of the texture, whereas matches to square-wave patterns stress the distribution of the luminance transitions in a cortical mechanism.

It is of course possible that the higher spatial frequency (cortical) channels are constructed from lower level channels that are much broader. If there are four low-level, broad channels, then a pair-wise comparison of the outputs of these low level channels

would yield six high frequency channels of narrower band width ($3+2+1$) with the peak sensitivity of the lowest high frequency channel near 3 c/deg. This lower limit agrees with that found by Blakemore and Campbell (1969).

Finally, returning to the usefulness of narrow-band, high frequency channels, one cortical structure that appears ideally situated and constructed to process luminance (or contrast) steps is the cortical column. Recently, Maffei and Fiorentini (1977) and also Pollen (1977) have reported that each cortical column in the cat contains at least two spatial frequency components separated by about an octave or slightly more. If the separation included a factor of 3, then, depending upon the proportions of each spatial frequency, the column as an entity itself could be more responsive to a square-wave or edge than to a pure sinusoid.*

iii) Estimate of Physiologic Primaries

Just as in color mixture one naturally asks which set of primaries is physiologically unique, the same question may be asked of texture matching functions. Namely, which matching functions describe the filters used by the human visual system? Certainly there are a very large number of transformations of texture matching functions given in Table II (no eye movements) that would be possibilities. Even if one requires that such fundamental texture matching functions be non-negative, there are still an unlimited number of transformations of these functions. Thus, the experimenter needs techniques for imposing further constraints upon the transformations of these texture matching functions.

*Recently, Marr and Poggio (1978) have suggested the importance of zero crossings in pattern analysis. These results support this notion and together with Maffei's result, suggest the cortical column as the basic neural unit of analysis.

In color one such set of constraints came from observers who were color blind and thus needed only a reduced set of primaries in order to make all matches. Thus, we can search for individuals who are texture blind and will need only three or perhaps only two spatial frequencies in order to match all possible textures. To date, however, such a search has been unproductive and other methods may be more profitably used for initial determinations.

For example, Spitzberg and Richards (1975) have shown that if the sine-wave gratings are flashed only briefly, then one portion of the frequency spectrum is attenuated more than the other. If the threshold for detecting a 50 msec modulation of a sine-wave is compared with the detection of a continuously modulated sine-wave then the spatial frequencies in the neighborhood of 10 to 12 c/deg become much harder to detect than any other spatial frequency. By comparing the threshold modulation for pulsed and continuously presented sine-waves, these authors discovered a function of spatial frequency that had the characteristics of a high frequency filter. This filter with a peak sensitivity near 10 c/deg may have a physiologic basis because suitable transformations of the texture matching function can yield a function resembling this filter (Richards and Polit, 1974).

A second manipulation that yields a spatial filter located near 1 c/deg is to compare thresholds for detecting a vertical grating only 0.2 deg high with thresholds for detecting the same grating 2 deg high (Spitzberg, 1975). Again, this filter is closely matched by a suitable transformation of the texture matching primaries (Richards and Polit, 1974).

Note that in each case the location of peak sensitivity of the measured spatial filter lies near a spatial frequency of one of the primaries of Table II (no eye movements). This correspondence is not coincidental, for a considerable amount of searching was conducted during pilot studies to determine the "optimal" primaries that yielded the best matches. We expect, therefore, that the location of the peak sensitivity of the physiologic fundamentals will lie near 1, 3, 6 and 11 c/deg.

A third possible method for uncovering the fundamental spatial filters is to compare texture matches in the fovea with those viewed eccentrically in the periphery. Again, drawing upon analogies with color vision, the spectral response characteristics of the rod-free fovea are quite different from those of the rod-dominated periphery. By suitable comparisons between foveal and peripheral viewing it is possible to isolate one or the other mechanism.

A comparison of Tables II (central) and III (extra-foveal) provide some data bearing upon this method. Note that the primaries are the same for both retinal positions. This suggests, just as it does in color vision, that the underlying spatial filters are the same at both the central and the eccentric retinal position. This finding is counter-intuitive, as we have been led to believe that the "grain" of visual processing becomes coarser the farther we move from the fovea. The trend toward increased sensitivity at lower spatial frequencies as field size increases suggests this coarsening of the visual metric, as do the ganglion cell counts.

However, such a coarsening for higher-level texture analysis may in fact be unrealistic. First, consider the problem of analyzing the textural gradients of a curved vertical cylinder (or sphere) with fixation at the center of the surface face. Clearly, the highest spatial frequency components are now outside the fovea, whereas the lowest are at the fovea itself. Similarly for the luminance gradients along a cylinder, where the shallowest gradient is at fixation and the steepest eccentric to the fovea.

For surface shape, therefore, an isotropic and homogeneous representation of spatial frequencies may be a more powerful analytical weapon than the non-isotropic representation.

Furthermore, eye movements of 6'-30' would be very suitable for driving the higher spatial frequency analyzers outside the fovea, even if their relative numbers decreased. (See Greenwood, 1972, for the effect of eye movements on spatial frequency sensitivity.)

The fact that the envelope of the threshold sensitivity function shifts to lower spatial frequencies with eccentricity no way obviates the possibility that the underlying spatial filters remain roughly the same. Shallow gradients and coarse spatial frequencies provide very useful information about surface structure in the region of fixation, just as high frequency information is important outside the fovea.

Why, then, does the visual system have a high ganglion cell density in the fovea? Two reasons come to mind: 1) the representation of stereopsis in the sagittal plane would require a nasal-temporal overlap and consequently lead to higher neural counts either at the retina or cortex if acuity is not to be sacrificed; 2) the multiple visual pathways leaving the retina would also lead to a higher ganglion cell density for central vision if each pathway is to include a representation of central (but not the same degree of peripheral) vision. This latter view would yield a visual system constructed like a stack of discs of varying diameters, with each disc encoding a separate visual function. (See also Koenderink in: Spekrijse and van der Tweel, 1977).

VI. GENERALIZED COLORIMETRY

Colorimetry is a method used to describe spectral sources that will appear identical to a representative human observer. Its success lies in the fact that color perception in man is based upon only three different types of "filters," specifically the absorption spectra of three different receptors. Whenever these three types of receptors are equally innervated by two physically different spectral lights, then these two lights will be indistinguishable perceptually.

Colorimetry is a one-to-one mapping of the relative absorptions of the three receptor types. The empirical observation that only three variables are necessary to characterize all possible color equivalences is a demonstration that color perception is based upon the sampling of only three (overlapping) regions of the visible spectrum.

The power of colorimetry lies not only in its practical applications to color rendition and reproduction, but also includes advances in our understanding of wavelength processing by man. In particular, when the nature of color blindness was first quantified using colorimetric methods, then the deficit could be represented precisely as a two variable system as opposed to the normal three. Given some simple assumptions about the nature of the phototransduction process, these reduced systems of the color-blind observer could be used to estimate the absorption characteristics of the normal human pigments (Helmholtz, 1890). The methods of colorimetry thus not only can identify the number of filters used by man to sample a continuum, such as wavelength, but also can characterize their properties if reduced cases are compared. How can this powerful, psychophysical method be adopted for more general use?

1. PRINCIPAL FEATURES OF COLORIMETRIC APPROACH

To proceed to develop a generalized colorimetry, we first point out four important constraints upon the method. These constraints deal in part with the concept of

"color-matching functions" that are the primary measurements of colorimetry. In color analysis, a matching function shows the amount (radiance) of a fixed wavelength "primary" that is needed to create a "match" to any arbitrary test wavelength of 1 unit strength. Thus, at each test wavelength, the value of the matching function shows the contribution of the "primary" wavelength to a match that will look like the test wavelength. Three such matching functions are needed to specify how all possible wavelengths may be "matched" by adding together the fixed primaries in the appropriate amounts.

An important feature of the colorimetric approach is that once the match to any wavelength is specified, then any spectral source can be matched merely by adding together the matches to its wavelength components. By the same token, the set of primaries can be changed by the appropriate addition or subtraction of the original set of matching functions. The underlying receptor sensitivities represent one such linear transformation.

For the success of the colorimetric approach, therefore, we may identify the following constraints:

- i) Equivalence Dimension:
Wavelength is a suitable dimension along which matching functions can be measured.
- ii) Uniqueness Property:
Stable and unique filters or "channels" are present (i.e. the different cone pigments).
- iii) Linearity Property:
Alternate sets of matching functions can be derived by adding or subtracting the members of the original set.
- iv) Intensive Property:
Color is an intensive variable that does not depend upon extent, hence spatial factors may be ignored.

For the success of the Generalized Colorimetric Method, a dimension for constructing matches or Equivalences (i) must be available and (ii) the sensory attribute to be studied must have filters or channels that sample this dimension uniquely. As we shall see, it is not necessary that Linearity (iii) hold exactly nor for the sensory variable to be Intensive (iv), although the interpretation is simpler and the analytical power greater when (iii) and (iv) are also valid.

Given the above assumptions, the generalized colorimetric technique proceeds in two steps: First, the minimum number of narrow-band stimuli necessary to create an equivalence to a broad-band distribution is determined. This is analogous to finding the minimum number of wavelengths needed to "match" a "white". Next, the matching functions are measured.

2. GENERALIZED COMPLEMENTS

Complementary lights are pairs of different spectral sources which mixed together will produce a "white". Because color is a three-variable system, complementary pairs of stimuli can be found that appear identical to a broad-band stimulus. The fact that many such pairs can be found demonstrates, given our assumptions above, that the human color processing is based upon no more than three different "filtered" samples of the wavelength dimension.

To show the relation between the number of complements and the underlying response or matching functions, refer to Fig. 6.1. In the top illustration, the sensitivities of two filters or response functions are shown along an arbitrary Equivalence dimension characterized by a horizontal line. A broad-band stimulus with a flat distribution along this dimension would innervate both response functions equally. But a narrow-band stimulus located at the intersection of the two sensitivity distributions (arrow) will also activate each response function equally. Hence for two independent filters or response functions,

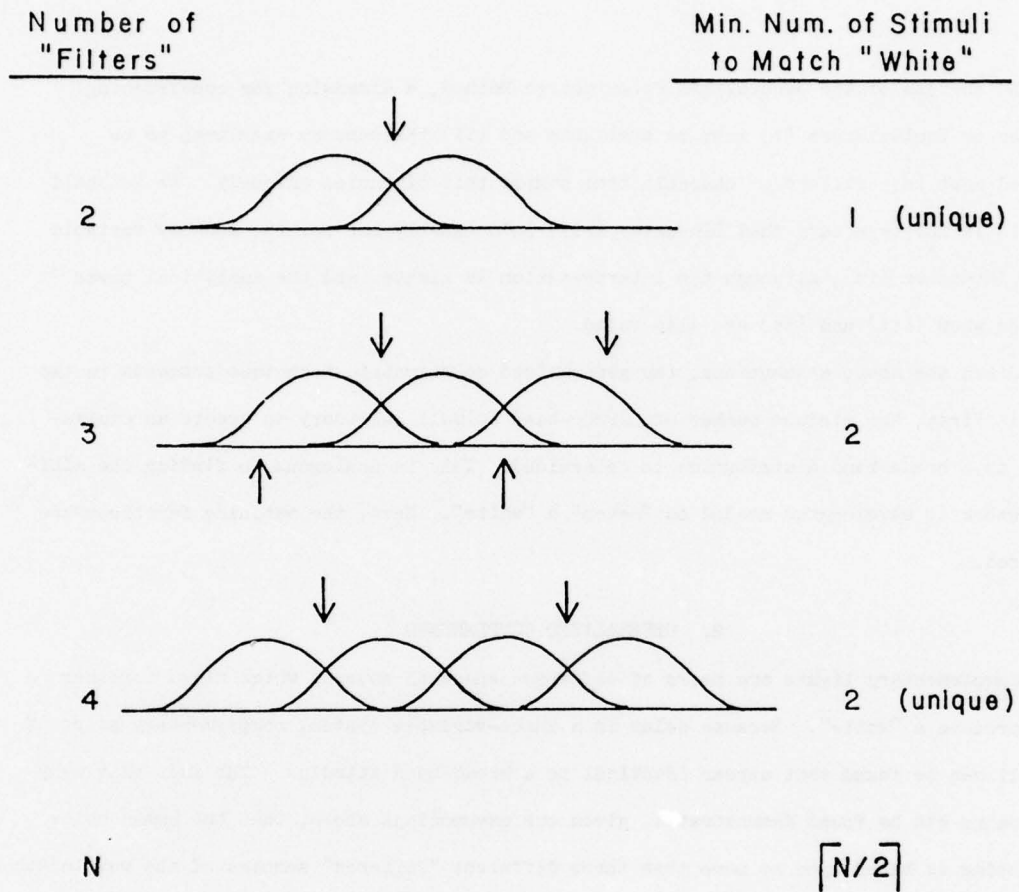


Figure 6.1

Hypothetical sensitivity distributions of channels along a stimulus dimension schematized by the horizontal line. The positions of narrow-band stimuli are indicated by arrows.

only one narrow-band stimulus is needed to create an equivalent sensation, and this choice is unique for a given "white".

Clearly, even if the areas under each response function were unequal, a "match" could still be found between a flat broad-band source and a simple narrow-band stimulus. The position of the narrow-band stimulus need only be moved toward the side of the function having the least area so that the ratio of the vertical line intercepts of the two functions equals the ratio of the convolutions of the source with the two response functions. The strength of the narrow-band stimulus can then be adjusted appropriately. In a similar manner, any arbitrary broad-band source can be shown to be "matched" by a single, unique narrow-band stimulus, regardless of the nature of the waveform of the "white."

It is also not necessary that the filters or response functions have unimodal distributions for a unique solution. However other narrow-band stimuli might be found to match certain broad-band sources under two circumstances:

- a) The continuum or Equivalence Dimension is closed (such as if it were a circular locus), or
- b) the matching narrow-band stimulus lies between the modes of one response function.

For the top illustration, a closed continuum would always lead to two possible solutions if both ends of each response function overlapped. At present, to simplify the preliminary analysis we will assume that the Equivalence Dimensions are not closed and that the response functions are unimodal.

To formalize the first result in Fig. 6.1 let the response functions be symbolized as $R_i(\lambda)$ and the narrow-band stimuli be designated as $S_j(\lambda)$. For an arbitrary energy source, $E(\lambda)$, we have

$$\begin{aligned} \int_{\lambda} E(\lambda) * R_1(\lambda) &= C_1 \\ \int_{\lambda} E(\lambda) * R_2(\lambda) &= C_2 \end{aligned} \tag{1}$$

where C_i is the output of the response function or "channel" activity. Let the ratio of activity of the two channels be

$$C_1 : C_2 = k \quad (2)$$

Now choose $S(\mu)$ such that $R_1(\mu) : R_2(\mu) = k$. Such a μ can be found because all possible ratios of R_1/R_2 exist since $R_i(\lambda) \Rightarrow 0$ at the tails. Next find the amplitude, A , of $S(\mu)$ such that

$$A * R_1(\mu) = C_1 \quad (3)$$

Then substitution in equation (2) shows that

$$A * R_2(\mu) = C_2 \quad (4)$$

and hence $A * S(\mu)$ will match $E(\lambda)$ since each elicit the same responses in R_i . [Note that we do not require linearity in the scaling of $R_2(\mu)$ for this solution. If nonlinearities in the amplitude scaling occur, then these can be offset by choosing another $S(\mu)$.]

Consider next the case where three response functions are used to sample a continuum, as in color vision. Here, as shown by the second illustration in Fig. 6.1, many pairs of narrow band stimuli can be found that stimulate both functions equally (only the two most obvious pairs are shown). However, although the number of complementary pairs is unlimited, the range over which they may occur is not. For example, as an extreme left-most (lower) stimulus encroaches more and more into the middle response function, the lower-right arrow must move to the right to reduce its stimulation of the same middle response function, until finally the lower pair of arrows will match the position of the upper pair. But the opposite argument applies to the upper pair of arrows, which must move to the left. Hence, stimuli lying in the central portion of the middle response function have no complements, unless the Equivalence dimension is closed.

To show more explicitly that a stimulus $S_1(\mu)$ has a complement as long as $R_1(\mu) > R_2(\mu)$, let $S_2(\lambda)$ innervate $R_3(\lambda)$ and $R_2(\lambda)$ such that $R_3(\lambda) > R_2(\lambda)$. In particular, at μ and λ , let

$$\begin{aligned} R_2(\mu) : R_1(\mu) &= k_1 \\ R_2(\lambda) : R_3(\lambda) &= k_3. \end{aligned} \tag{5}$$

We then wish to show that

$$S_1(\mu) + A \cdot S_2(\lambda)$$

yields

$$R_1 = R_2 = R_3. \tag{6}$$

If the amount, A , of S_2 is such that

$$A \cdot R_3(\lambda) = R_1(\mu) \tag{7}$$

then

$$A \cdot R_2(\lambda) = A \cdot k_3 \cdot R_3(\lambda) = k_3 \cdot R_1(\mu). \tag{8}$$

But

$$R_2(\mu) = k_1 \cdot R_1(\mu). \tag{9}$$

Hence the total amount of R_2 is the sum of (8) and (9) to yield

$$R_2(\mu+\lambda) = (k_1+k_3) \cdot R_1(\mu). \tag{10}$$

Thus,

$$k_1+k_3 = 1. \tag{11}$$

In practice, given an $S_1(\mu)$ with a known R_2/R_1 ratio, $S_2(\lambda)$ can be found by equation (11) and then A chosen to satisfy (7).

Once again, the outputs of the response functions do not have to be linear functions of the inputs in order for complementary solutions to be found. However, equation (11)

may become invalid if non-linearities are present.

The last and lowermost illustration in Fig. 6.1 shows the case where the continuum is sampled by four response functions. In this case, only one solution for narrow-band complements are created at the intersections of the two left and two right-most response functions, at least for a flat broad-band "white".

From Fig. 6.1 it should now be clear that whenever an even number N of response functions sample a continuum that is not closed, then the minimum number of narrow-band stimuli needed to match a broad-band "white" will be $N/2$. The solution will be unique with the stimuli located at the intersections of pairs of response functions.

When the number N of response functions is odd, however, the minimum number of narrow-band stimuli will be the integer value of $N/2$. For example, in the case of five response functions, three narrow-band stimuli will be required. In terms of Fig. 6.1 the solution for the five channel case may be visualized better either as the solution for two pairs of functions plus one [where $S_j(\lambda)$ can be at an isolated tail of R_j], or as one pair of functions plus three. Note that the solutions for an odd number of response channels will not be unique, thus distinguishing the even and odd cases where $\lceil N/2 \rceil$ is equal.

It now should be clear that by appropriate pairing of the response functions, that complementary narrow-band stimuli can always be found. In the case where the number of response functions is even, merely pair the first two response functions and apply equation (3). Then proceed to the next two and repeat the procedure, etc. For an odd number of response functions, either treat the last, unpaired response function in isolation by stimulating its "tail", or determine the solution for the last triplet by using equation (11). Note that although the use of the equations may require linearity in the input-output relations of the underlying "channels", solutions can still be found by iterative trial and error even if these relations are non-linear. The Linearity Property becomes important only if transformations between matching functions are to be made.

To summarize, for an open Equivalence dimension sampled by N response functions or "channels", the minimum number of narrow-band stimuli matching a broad-band "white" will be the integer value of $N/2$. If the solution does not require unique (in the sense of highly restricted) narrow-band stimuli, then the minimum number of sensory filters sampling the continuum is not greater than twice the number of matching narrow-band stimuli, less one.

3. EXAMPLES OF GENERALIZED COMPLEMENT TECHNIQUE

Texture Metamers: Earlier, Section II described texture matches created from vertical bars of equal width but having randomly chosen gray levels (0-63). Such a pattern is a band of "white noise". These "noise" patterns containing 64 gray levels could be matched by a simplified texture created from bars of the same width, but having only one of three gray levels. The exact choice of grays were not important (see Figs. 2.9, 2.10 and 6.2). Thus, the human visual system must be "filtering" this kind of noisy texture information. Along a gray scale continuum, the Generalized Colorimetry approach would suggest that only 5 gray level response functions at most are required to characterize these matches. (In fact, probably only three are being used, with the extra gray being required to create a more appropriate spatial frequency match.)

Furthermore, all such "white noise" textures, regardless of whether or not they are created from bars of equal width or not, can be matched by only three spatial frequencies (see Fig. 2.2 for example). The exact choice of spatial frequencies is not important (although the waveform may be). Once again, the Generalized Colorimetric analysis would suggest that at most 5 spatial frequency response "channels" are used in this type of analysis. (Without eye movements, Section III demonstrates that in fact only 4 spatial frequency response functions are present.)



Figure 6.2

The generalized colorimetric method applied to gray level encoding. The gray level of most of the checks in the pattern have one of 63 levels, chosen randomly. Which half has only three gray levels represented? (Courtesy of M.D. Riley)

Orientation Discrimination: As another example of the power and generality of the method, consider the question "How many orientation 'channels' participate in the global perception of textural structure?" The dimension along which these channels can be represented is obviously closed, with values 0 to 180 degrees for simple line elements. A suitable broad-band stimulus is merely a texture constructed from line elements having random orientations (left portion of Fig. 6.3a). A narrow-band stimulus is represented by line segments at a single orientation. As shown by the right portion of Fig. 6.3a, two such narrow-band stimuli are unlikely ever to yield a match to the "white" regardless of the choice of orientations or relative density of the line elements. Fig. 6.3b compares a broad-band "white" texture with one constructed from elements of three orientations. A match is still not possible with this choice of orientations and density. However, a slight increase in density of the right pair might yield a successful match if the patterns are not scrutinized.

Fig. 6.3c illustrates that four orientations can yield a successful match, and experiment has shown that many different quadruples of orientations will work (M.D. Riley, 1977). Thus, we suspect that for any given broad-band pattern of orientations, there will be only one set of triplets that will yield an equivalence. Our preliminary estimate of the number of channels used for this task is therefore six. This result would suggest a "channel" width of about 30° --a value consistent with other psychophysical studies (Campbell and Levinson, 1972).

Flicker Discrimination: White noise temporal flicker can easily be matched by an appropriate combination of two sinusoidally flickering lights. Depending upon the sampling time, either many solutions are possible if the time is short, or few if the time is long. This result would imply that three or four flicker "channels" are used to sample the temporal frequency spectrum. We shall see shortly that three response functions yield good flicker matches under most conditions (Richards, 1975).

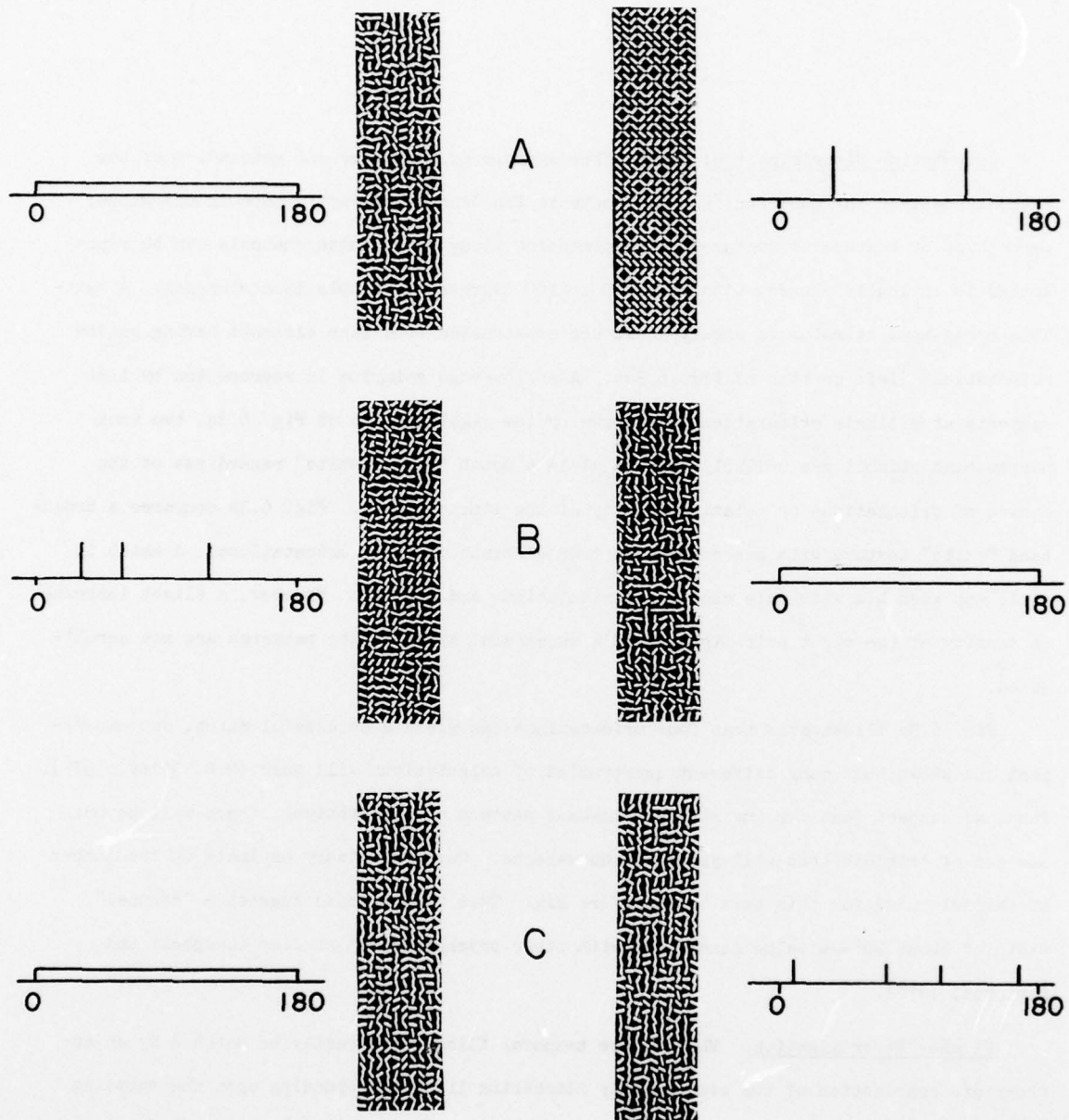


Figure 6.3

The generalized colorimetric technique applied to the orientation of texture elements. Each panel from top to bottom has, respectively two, three, and four orientations. (Adapted from M. Riley, 1977 by A. Witkin.)

4. DERIVATION OF MATCHING FUNCTIONS

Once again, we will follow the guidelines of colorimetry and determine first a set of "matching" (distribution) functions that are isomorphic with the underlying response functions or "channels". These matching functions describe the amounts of fixed (narrow-band) primaries that are required to create metameric matches to (narrow-band) stimuli located anywhere on the continuum of interest. An upper bound on the number of primaries required is set first by the matches to "white noise".

A trial and error procedure is then usually used to determine the optimal choice and minimum number of primaries. Furthermore, the linearity assumption must be invoked, permitting the "desaturation" of any narrow-band test stimulus by one or more of the primaries. When a stimulus is desaturated, by a primary, the amount of the primary used is given a negative sign in the equivalence relation. Thus,

$$S_i(\lambda_i) \equiv a_i * P_1(\lambda_1) + b_i * P_2(\lambda_2) - c_i * P_3(\lambda_3)$$

indicates that to create an equivalence between the stimulus S_i located at λ_i and the two fixed primaries P_1 and P_2 , the third primary P_3 must be added in amount c_i to the test stimulus S_i .

If marked non-linearities are present, then the color matching technique will fail, as indicated by the discovery that the number of fixed primaries required to match all portions of the continuum will exceed the upper bound imposed by the "white noise" matches.

The presence of small, non-linearities, however, is no obstacle and may even become an asset. For example, if the primaries are chosen to be near the peaks of the underlying response functions, then the range of amplitudes of the dominant primary in this region will be smaller than if the "tail" of the response function is stimulated. In the last case, a considerable amount of the primary may be required to create a match to the test stimulus located near the peak sensitivity of that same response function, and the

non-linear effects will become more marked, as evidenced by excessive desaturants. As a general rule, therefore, primaries should be chosen to maximize the saturation of the test stimulus and to minimize the amounts of desaturants required.* With this strategy, the matching functions are more likely to reflect the sensitivity profiles of the underlying response functions. The more the desaturation required in the matches, the greater will be the dependence placed on the linearity assumption in order to transform the empirical matching functions to a representative set of all-positive response functions.

One of the more difficult problems in finding suitable primaries is to determine their spacing along the dimension of interest. This task can be simplified in two ways: First, the location of the "complementary" stimuli required to make "white noise" matches provide some cue to the location of optimal primaries. At least one and sometimes two of the best primaries will lie in between the adjacent complements, as can be seen by inspecting Fig. 6.1. A second method of simplification is to recognize that at any given primary, the amounts of all other primaries required for a match is zero. These zero crossings of the matching functions thus impose a constraint on the waveform of the matching functions. By assuming a waveform, a relation between the matching properties of those stimuli located located intermediate between any two primaries can be determined. For example, in the case where only two negative (desaturating) lobes are assumed, the size of the spacing between primaries (k) is related to the test stimulus T by the relation:

$$0.5(T) + a(k^{3/2}T) + b(k^{-3/2}T) \equiv c(k^{1/2}T) + d(k^{-1/2}T) \quad (12)$$

for four underlying response functions sampled on a logarithmic scale. If suitable

*If the system is linear, matches can be made to either a maximally saturated field (Wright, 1928) or to a minimally saturated "white" field (Guild, 1931) and the results will be identical. In other words, one set of functions can be obtained by a linear transformation of the other.

coefficients a , b , c and d can not be found, then k is too large. This relation was used previously to determine the maximum spacing possible between spatial frequency primaries used for texture matches. (See Section II.4.)

5. EXAMPLES OF DERIVATION OF RESPONSE FUNCTIONS

i) Spatial Frequency Matches:

The first two sections of this report document how texture matches were found using the spatial frequency continuum. Figure 2.5 summarizes the results for the case where eye movements were minimal. In order to determine the underlying response functions or "channels" from those matching functions, linearity must be assumed, and other constraints must be imposed. One obvious constraint is that the response functions be non-negative everywhere. Other constraints and a possible solution are given in Richards and Polit (1974).

ii) Temporal Frequency Matches:

Fig. 6.4 summarizes preliminary flicker matching functions showing the amounts of three primary flicker frequencies needed to match any arbitrary sinusoidal flicker in the range 0.5 to 30 c/deg (Richards, 1975). Note that little desaturation is required in most cases, and thus these matching functions probably come close to representing the underlying flicker "channels". One striking characteristic not obvious until these matching functions were measured is that all high frequencies above 12 c/sec can be made to look the same by an appropriate adjustment of contrast. (At least for this particular matching situation where fixation was held between two panels 3° wide x 2° high and separated by $2/3^\circ$.)

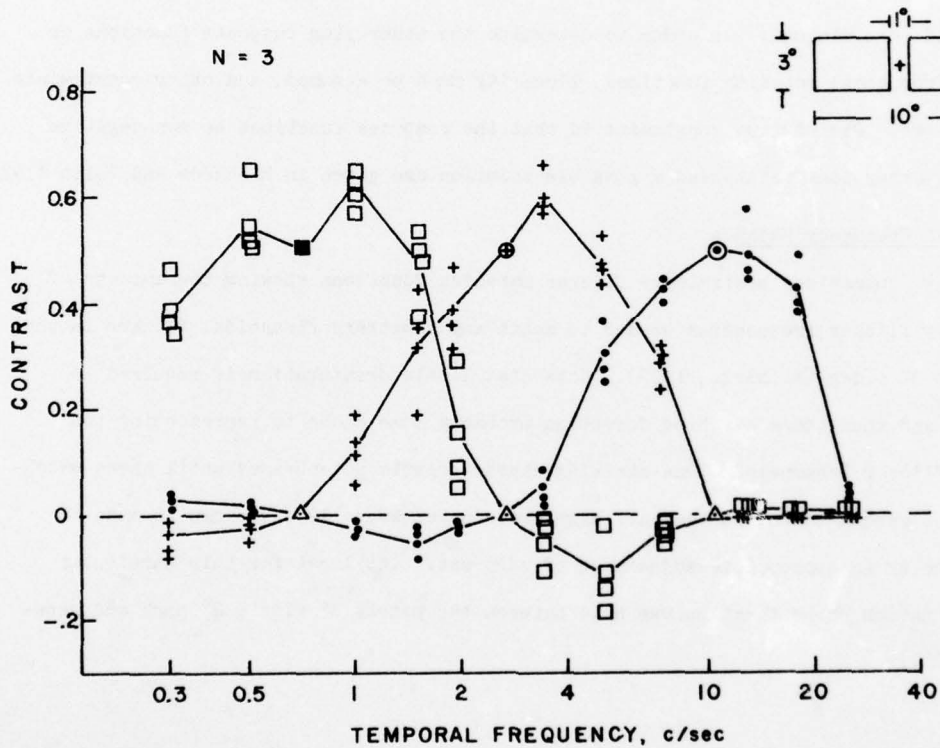


Figure 6.4

Flicker matching functions for three observers (the fourth point is the mean value), using primaries of 10.2, 2.7 and 0.7 Hz.

6. INTENSIVE VERSUS EXTENSIVE VARIABLES

The above two examples differ in that one involves intensive matching (flicker) whereas the other is an extensive variable (spatial frequency). In the case where extensive variables are used, the spatial extent of the pattern must clearly influence the result. For example, in the spatial frequency texture matches, if only a $1/2$ deg field were used, the number of primaries required would be reduced to probably two. At the opposite extreme, how can we insure that as the field is enlarged, the number of required primaries will not increase? In fact, they might, although this is not the case for texture matches made with minimal eye movements, for the same set of primaries suffice in the neighborhood of the fovea and at 6° retinal eccentricity (see section II).

When eye movements are allowed, however, up to six primaries are needed over the range of 0.2 to 30 c/deg, and it is expected that others would need to be added as the field width is further enlarged. Nevertheless, for any given spatial frequency, four primaries will suffice. This can be seen by inspecting Table I, where all primaries but four will have zero magnitudes at any given test frequency.* Thus, although this result may be misleading for an interpretation of the nature of the response functions at any given retinal eccentricity, it does demonstrate a technique for overcoming difficulties that may be imposed by spatial variables when extensive matching functions are to be determined.

*Because of the correlation between decreasing spatial frequency and increasing retinal eccentricity, one might conclude that at any retinal eccentricity, four spatial filters will suffice for spatial frequency analysis and that the coarseness of these four increase with eccentricity. Such a conclusion must be viewed with caution, however, because the presence of eye movements adds temporal cues to the detection task (Koenderink, 1972; Smith, 1977; Greenwood, 1972.)

7. HIERARCHICAL PROCESSING

All sensory systems are constructed in a hierarchical fashion, where the outputs of one level serve as the input to the next. Although Generalized Colorimetry is measuring response functions along one physical continuum, there is an implicit relation between this continuum and the sensory dimension that serves as a basis for setting equivalences. These implicit relations can be further extended by an assumption that a given sensory dimension corresponds to one level of processing. Although probably false, this assumption has heuristic value.

For example, in texture perception, the spatial frequency matches made using sinusoidal primaries yield one set of matching functions that reflect constraints imposed by subcortical "channels". But at the next level in the cortex, the appropriate matching functions change to emphasize edges and sharp lines, requiring a new set of primaries with a different waveform. The new set of "square-wave" matching functions can be interpreted as an (differencing) operation imposed upon the activities of the broad-band "channels" that serve as their inputs. Such functions would serve as a more appropriate basis for computing the distribution of steps in luminance or contrast.

Now if the (differencing) operation were completely linear, then both sets of matching functions would be related to one another by a linear transformation. Furthermore, the same sensory dimension would be used to determine equivalences for both sets of matching functions. If either a new sensory dimension is used, or if the (differencing) operation is nonlinear, then the sets will be partially independent. But either possibility implies a new level of operation.

As another example of the relation between matching functions and levels of processing, consider the equivalences between textures containing identical elements of different orientation (Fig. 6.3). Six orientation "channels" are implied by the matches to textures

with random orientations. This limitation in orientation processing may be a grouping constraint imposed upon the linking of "points" in an image (i.e. as would be the case if the mosaic has hexagonal packing with horizontal and vertical asymmetry; Richards, 1970). If the limitation occurs in this manner, then any grouping operation performed at the next higher level is constrained to use only these six selected orientations, and this new level would have associated with it a new sensory dimension other than orientation. One such dimension, for example, could be related to the junctions of Guzman (1968) and Waltz (1975).

On the other hand, the orientation limitation could be directly imposed by the junction analysis itself (Sakitt, 1977), which would be performed presumably upon a population of feature detectors (such as the simple cells of Hubel and Wiesel, 1959, 1968) that have a wide spectrum of distributed orientations.

The fact that two interpretations are possible, however, in no way detracts from the power of the Generalized Colorimetry approach. If the process of junction analysis has imposed the limitation upon the discrimination of texture orientation, another dimension of attributes should be sought to analyze processing at the level of the "simple cell" feature detector. Similarly, the first interpretation implying less specific grouping or "packing" constraints suggests a search for a sensory dimension that would reflect junction analysis. In either case, the method of Generalized Colorimetry helps to dissect and explore the various levels of hierarchical processing.

VII. MAJOR CONCLUSIONS

APPLIED

1. For most practical purposes, the parallel linear structure of texture can be simulated by a suitable weighted combination of four fixed spatial frequencies.
2. For "fuzzy" or "blurred" textures, the fixed spatial frequencies should have a sinusoidal waveform. For textures containing mostly edges or step changes in contrast, a square waveform and a different set of higher, more closely spaced spatial frequencies should be used.
3. The vertical and horizontal linear structure of a texture act independently but the same set of primaries suffice for texture matches at both orientations (retinal coordinates), but not at oblique (45 deg) orientations. For oblique orientations, the primary spatial frequencies should be scaled down by 0.7x.
4. Two dimensional textures with components at any orientation can probably be simulated in most cases by a suitable combination of one dimensional textures each having one of only four but no more than six orientations.
5. A visual graphics display system is described that has wide experimental usefulness and capabilities for visual psychophysics.

BASIC

1. Central texture vision for "blurred" patterns probably utilizes only four low level spatial filters or "channels" at any given orientation.
2. At 45 deg retinal orientation, the spatial frequency sensitivity of the "channels" is displaced by .7x to lower frequencies as compared with vertical or horizontal orientations.

3. These "low level" spatial filters probably precede the "cortical" channels revealed by adaptation studies.
4. These four low level spatial filters do not change size with retinal eccentricity within the central 10° of vision. Their relative proportions and contributions to the sine-wave sensitivity functions do change, however.
5. A weighted sum of the texture matching functions may be used to reconstruct the sine-wave sensitivity function (Additivity Property).
6. The texture matching functions behave reasonably linearly for small changes in the primaries, permitting transformations from one set of primaries to another. (Linearity Property). However, if high-frequency primaries are used, linearity fails.
7. Estimated physiologic response functions are described within the limits of accuracy of the linearity property.
8. Textures containing many step changes in contrast are probably analyzed by "higher-level" (cortical) spatial filters with a narrower band-width. The spatial distribution of these step changes is more important than the magnitude of the contrast change.
9. A new powerful psychophysical technique is described, called "Generalized Colorimetry" that in principal can be used to analyze any level of sensory processing, including modalities other than vision.

VIII. SIGNIFICANCE

Texture is one of the primary properties of an object. Like color, texture is a quality that helps the human observer to define and identify objects. Yet we know very little about human texture perception. What is its basis? How good are we at identifying textures? Are we as good as a Fourier pattern analyzer? At present, the most important aspect of the research suggests that texture analysis of patterns with only one orientation component is performed by only four "filters". Thus all one-dimensional textures may be completely specified in terms of only four primaries, at least over the range of focal vision. Such a specification will describe all equivalences between textures that are constructed from similar basis (i.e. such as sine-wave or square-wave gratings or probably even line elements). For two-dimensional textures, four such sets must be considered, one set for each of four orientations. This limitation on man's ability to analyze textures is a non-trivial fact with important practical consequences. In the domain of color perception, if it were necessary to describe all colors in terms of the precise wavelength composition, then the transmission of chromatic information would not have become a feasible possibility. The fact that the human observer analyzes textures on the basis of only a few filters, then a considerable saving in the transmission of texture information may be gained. This practical benefit far outweighs, but in no way diminishes the further gains that we will achieve in our understanding of the human visual system.

IX. Bibliography

- Abadi, R.V. and Kulikowski, J.J. (1973) Linear summation of spatial harmonics in human vision. Vis. Res. 13, 1625-1628.
- Appel, K. and Haken, W. (1977) The solution of the four-color-map problem. Sci. Am. 237, 108-121 (Oct.).
- Atkinson, J. and Campbell, F.W. (1974) The effect of phase on the perception of compound gratings. Vis. Res. 14, 159-162.
- Blakemore, C. (1970) A new kind of stereoscopic vision. Vis. Res. 10, 1181-1199.
- Blakemore, C. and Campbell, F.W. (1969) On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images. J. Physiol. 203, 237-260.
- Brown, P.K. and Wald, G. (1964) Visual pigments in single rods and cones of the human retina. Science 144, 45-52.
- Bryngdahl, O. (1966) Perceived contrast variation with eccentricity of spatial sine-wave stimuli. Vis. Res. 6, 553-565.
- Burton, G.J. (1976) Visual detection of patterns periodic in two dimensions. Vis. Res. 16, 991-998.
- Campbell, F.W. and Kulikowski, J.J. (1966) Orientational selectivity of the human visual system. J. Physiol. 187, 437-445.
- Campbell, F.W. and Robson, J.G. (1968) Application of Fourier analysis to the visibility of gratings. J. Physiol. 197, 551-566.
- Carlson, C.R.; Cohen, R.W. and Gorog, I. (1977) Visual processing of simple two-dimensional sine-wave luminance gratings. Vis. Res. 17, 351-358.
- Cornsweet, T.N. (1970) Visual Perception. Academic Press, New York.
- Cowan, J.D. (1977) Some remarks on channel bandwidths. Neurosciences Res. Prog. Bull. 15, 492-517.
- Davidson, M.L. (1968) Perturbation approach to spatial brightness interaction in human vision. J. Opt. Soc. Am. 58, 1300-1309.

- DeLange, H. (1958) Research into the dynamic nature of the human fovea-cortex systems with intermittent and modulated light. J. Opt. Soc. Am. 48, 777-789.
- DePalma, J.J. and Lowry, E.M. (1962) Sine-wave response of the visual system. J. Opt. Soc. Am. 52, 328-335.
- DeValois, K.K. (1977) Independence of black and white: phase-specific adaptation. Vis. Res. 17, 209-216.
- DeValois, K.K. (1977) Spatial frequency adaptation can enhance contrast sensitivity. Vis. Res. 17, 1057-1066.
- Doorn, van A.J.; Koenderink, J.J. and Bonman, M.A. (1972). The influence of retinal inhomogeneity on the perception of spatial patterns. Kybernetik 10, 223-230.
- Fiorentini, A.; Siretencanu, R. and Spinelli, D. (1976) Lines and gratings: different interocular aftereffects. Vis. Res. 16, 1303-1309.
- Flock, H.R. and Moscatelli, A. (1964) Variables of surface texture and accuracy of space perceptions. Percept. and Motor Skills 19, 327-334.
- Franzen, O. and Berkley, M. (1975) Apparent contrast as a function of modulation depth and spatial frequency. Vis. Res. 15, 655-660.
- Gibson, J.J. (1950) The Perception of the Visual World. Houghton Mifflin Co., Boston.
- Gilinsky, A.S. (1968) Orientation-specific effects of patterns of adapting light on visual acuity. J. Opt. Soc. Am. 58, 13-18.
- Glass, L. and Perez, R. (1973) Perception of random dot interference patterns. Nature 246, 360-362.
- Graham, N. (1972) Spatial frequency channels in the human visual system. Vis. Res. 12, 53-68.
- Graham, N. (1977) Visual detection of aperiodic spatial stimuli by probability summation among narrow band channels. Vis. Res. 17, 637-652.
- Graham, N. and Nachmias, J. (1971) Detection of grating patterns containing two spatial frequencies. Vis. Res. 11, 251-259.
- Graham, N. and Rogowitz, B.E. (1976) Spatial pooling properties deduced from the detectability of FM and quasi-AM gratings; a re-analysis. Vis. Res. 16, 1021-26.

- Green, B.F. Jr.; Wolf, A.K. and White, B.W. (1959) The detection of statistically defined patterns in a matrix of dots. Am. J. Psychol. 72, 503-520.
- Greenwood, R.E. (1972) Some effects of involuntary eye movements. J. Opt. Soc. Am. 62, 101-103.
- Gruber, H.E. and Clark, W.C. (1956) Perception of slanted surfaces. Percept. and Motor Skills, Monogr. Suppl. 2, 6, 97-106.
- Guild, J. (1931) The colorimetric properties of the spectrum. Phil. Trans. 230A, 149-187.
- Guzman, A. (1968) Computer recognition of three-dimensional objects in a visual scene. (Tech. Report AI-TR-228) MIT Artificial Intelligence Lab.
- Hamerly, J.R.; Quick, R.F. and Reichert, T.A. (1977) A study of grating contrast judgement. Vis. Res. 17, 201-207.
- Harvey, L.O., Jr. and Gervais, M.G. (1978) Visual texture perception and Fourier analysis. Perception and Psychophysics (in press).
- Helmholtz, H. von (1890) Treatise on Physiological Optics. J.P.C. Southall (Trans.) New York: Dover, 1962.
- Henning, G.B.; Hertz, B.G. and Broadbent, D.E. (1975) Some experiments bearing on the hypothesis that the visual system analyzes spatial patterns in independent bands of spatial frequency. Vis. Res. 15, 887-897.
- Hilz, R. and Cavonius, C. (1974) Functional organization of the peripheral retina: sensitivity to periodic stimuli. Vis. Res. 14, 1333-1337.
- Hubel, D.H. and Wiesel, T.N. (1968) Receptive fields and functional architecture of monkey striate cortex. J. Physiol. 195, 215-243.
- Jones, L.A. and Higgins, G.C. (1945) The relationship between the granularity and graininess of developed photographic materials. J. Opt. Soc. Am. 35, 435-457.
- Julesz, B. (1962) Visual pattern discrimination. IRE Trans. of the Prof. Group on Information Theory IT-8, No. 2, 84-92.
- Julesz, B. (1965) Texture and visual perception. Sci. Am. 212, 38-48 (Feb.).
- Julesz, B. (1971) Foundations of Cyclopean Perception. Univ. Chicago Press, Chicago.

- Julesz, B. (1972) Visual pattern discrimination. IRE Trans. on Information Theory IT-8, 84-92.
- Julesz, B. (1975) Experiments in the perception of visual texture. Sci. Am. 232, 34-43 (April).
- Julesz, B.; Frisch, H.S.; Gilbert, E.N. and Shepp, L.A. (1973) Inability of humans to discriminate between visual textures that agree in second-order statistics revisited. Perception 2, 391-405.
- Julesz, B. and Miller, J.E. (1975) Independent spatial-frequency tuned channels in binocular fusion and rivalry. Perception 4, 125-144.
- Kelly, D.H. (1966) Frequency doubling in visual responses. J. Opt. Soc. Am. 56, 1628-1633.
- Kelly, D.H. (1976) Pattern detection and the two-dimensional Fourier transform: flickering checkerboards and chromatic mechanisms. Vis. Res. 16, 277-287.
- Kelly, D.H. (1977) Visual contrast sensitivity. Optica Acta 24, 107-129.
- Kelly, D.H. and Magnuski, H.S. (1975) Pattern detection and the two-dimensional Fourier transform: circular targets. Vis. Res. 15, 911-915.
- Koenderink, J.J. (1972) Contrast enhancement and negative afterimage. J. Opt. Soc. Am. 62, 685-689.
- Koenderink, J.J. and Doorn, van A.J. (1974) Detectability of two-dimensional band limited noise. Vis. Res. 14, 515-518.
- Koffka, K. (1935) Principles of Gestalt Psychology. Harcourt Brace, New York.
- Kraft, A.L. and Winnick, W.A. (1967) The effect of pattern and texture gradient on slant and shape judgements. Perception and Psychophysics 2, 141-147.
- Kuffler, S.W. (1953) Discharge patterns and functional organization of mammalian retina. J. Neurophysiol. 16, 37-68.
- Kulikowski, J.J. (1976) Effective contrast constancy and linearity of contrast sensation. Vis. Res. 16, 1419-1431.
- Kulikowski, J.J. and King-Smith, P.E. (1973) Spatial arrangement of line, edge, and grating detectors revealed by subthreshold summation. Vis. Res. 13, 1455-1478.

- Limb, J.O. and Rubinstein, C.B. (1977) A model of threshold vision incorporating inhomogeneity of the visual field. Vis. Res. 17, 571-584.
- Maffei, L. and Fiorentini, A. (1971) Binocular depth perception with geometrical cues. Vis. Res. 11, 1299-1305.
- Maffei, L. and Fiorentini, A. (1977) Spatial frequency rows in the striate visual cortex. Vis. Res. 17, 257-264.
- Marks, W.B.; Dobelle, W.H. and MacNichol, E.F. (1964) Visual pigments of single primate cones. Science 143, 1181-1182.
- Marr, D. (1974) The computation of lightness by the primate retina. Vis. Res. 14, 1377-1388.
- Marr, D. (1976) Early processing of visual information. Phil. Trans. Roy. Soc. Lond. B 275, 483-524.
- Marr, D. (1977) Representing visual information. MIT Artificial Intelligence Lab. Report AIM 415.
- Marr, D. and Poggio, T. (1978) A theory of human stereo vision. (Submitted to Proc. Roy. Soc. Lond.)
- Maxwell, J.C. (1855) Experiments on color, as perceived by the eye with remarks on color-blindness. Trans. Roy. Soc. (Edinburgh) 21, 275.
- McBride, P. and Reed, J.B. (1952) The speed and accuracy of discriminating differences in number and texture density. Human Engr. Tech. Report SDC-131-1-3.
- Metzger, W. (1930) Reference in Gibson, J.J., pg. 5, The Perception of the Visual World. Houghton Mifflin Co., Boston, 1950.
- Minsky, M. and Papert, S. (1969) Perceptrons. MIT Press, Cambridge, Mass.
- Mitchell, D.E. and Ware, C. (1974) Interocular transfer of a visual after-effect in normal and stereoblind humans. J. Physiol. 236, 707-721.
- Mitchell, O.R. (1976) Effect of spatial frequency on the visibility of unstructured patterns. J. Opt. Soc. Am. 66, 327-338.

- Mostafavi, H. and Sakrison, D.J. (1976) Structure and properties of a single channel in the human visual system. Vis. Res. 16, 957-968.
- Pantle, A. (1973) Visual effects of sinusoidal components of complex gratings: independent or additive? Vis. Res. 13, 2195-2204.
- Pantle, A. and Sekular, R. (1968) Size detecting mechanisms in human vision. Science 162, 1146-1148.
- Pickett, R.M. (1962) Discrimination of constraint in random visual texture (Doctoral dissertation, Univ. Michigan, June, 1962). Ann Arbor Michigan, Univ. Microfilms, 1963, No. 63-427.
- Pickett, R.M. (1964) The perception of a visual texture. J. Exp. Psychol. 68, 13-20.
- Pickett, R.M. (1967) The perception of random visual texture. In: Models for the Perception of Speech and Visual Form. Ed. by W. Wathen-Dunn. MIT Press, Cambridge, Mass. pp. 224-232.
- Pickett, R.M. (1968) Perceiving visual texture: A literature survey. Aerospace Medical Research Laboratories Report AMRL-TR-68-12.
- Polit, A. (1976) Computer Generation of Textures. Behavior Research Methods and Instrumentation 8(4), 367-368.
- Pollack, I. (1972) Visual discrimination thresholds for one- and two- dimensional Markov spatial constraints. Perception and Psychophysics 12, 161-167.
- Pollack, I. (1973) Discrimination of third-order Markov constraints within visual displays. Perception and Psychophysics 13, 276-280.
- Pollen, D. (1977) Personal communication.
- Purks, S.R. and Richards, W. (1977) Visual texture discrimination using random-dot patterns. J. Opt. Soc. Am. 67, 765-771.
- Quick, R.F., Jr.; Hamerly, J.R. and Reichert, T.A. (1976) The absence of a measurable cortical band at low suprathreshold contrasts. Vis. Res. 16, 351-355.
- Quick, R.F., Jr. and Reichert, T.A. (1975) Spatial frequency selectivity in contrast detection. Vis. Res. 15, 637-643.

- Richards, W. (1971) The fortification illusions of migraines. Sci. Am. 224 88-96 (May).
- Richards, W. (1975) Flaxhback to Maxwell. J. Opt. Soc. Am. 65, 1200A.
- Richards, W. (1977) Stereopsis with and without monocular contours. Vis. Res. 17, 967-969.
- Richards, W. and Foley, J.M (1978) Superposition of tilted gratings. (Submitted to Vis. Res.)
- Richards, W. and Polit, A. (1974) Texture matching. Kybernetik 16, 155-162.
- Richards, W. and Purks, S.R. (1978) On random-dot discrimination. J. Opt. Soc. Am. (in press).
- Richards, W. and Quick, R.F. (1978) Two-dimensional spatio-temporal contrast-sensitivity functions. (Submitted to Vis. Res.)
- Richards, W. and Riley, M.D. (1977) Texture metamers. J. Opt. Soc. Am. 67, 1401A.
- Riley, M.D. (1977) Discrimination of bar textures with differing orientation and length distributions. MIT B.S. Thesis, Dept. of Elec. Engr. and Computer Science.
- Riley, M.D. (1977) Discrimination of textures with different orientation distributions. J. Opt. Soc. Amer. 67, 1401A.
- Robson, J.G. (1966) Spatial and temporal contrast-sensitivity functions of the visual system. J. Opt. Soc. Am. 56, 1141-1142.
- Rosenblatt, M. and Slepian, D. (1962) Nth order Markov chains with any set of N variables independent. J. Soc. Indus. and Appl. Math 10.
- Rosenfeld, A. (1967) Models for the perception of visual texture. In: Models for the Perception of Speech and Visual Form. Ed. by W. Wathen-Dunn, MIT Press, Cambridge, Mass. pp. 219-223.
- Sachs, M.B.; Nachmuis, J. and Robson, J.G. (1971) Spatial frequency channels in human vision. J. Opt. Soc. Am. 61, 1176-1186.
- Sakitt, B. (1977) Personal communication.
- Sansbury, R.V. (1977) Contrast movement detection of the square-wave third harmonic. Vis. Res. 17, 307-309.

- Santoro, T.P. and Fender, D.H. (1976) Rules for the perception of connectivity in random dot patterns. Vis. Res. 16, 973-981.
- Savoy, R.L. and McCann, J.J. (1975) Visibility of low spatial-frequency sine-wave targets: dependence on number of cycles. J. Opt. Soc. Am. 65, 343-350.
- Schnitzler, A.D. (1976) Theory of spatial-frequency filtering by the human visual system II. J. Opt. Soc. Am. 66, 617-625.
- Smith, R.A., Jr. (1977) Spatial-frequency adaptation and afterimages. Perception 6, 153-160.
- Spekreijse, H. and van der Tweel, L.H. (1977) Spatial Contrast. North-Holland Publishing Co.: Amsterdam.
- Spillmann, L. and Levine, J. (1971) Contrast enhancement in a Hermann grid with a variable figure-ground ratio. Exp. Br. Res. 13, 547-559.
- Spitzberg, R.M. (1973) On visual mechanisms and the processing of spatio-temporal information. MIT Ph.D. Thesis, Physics Dept.
- Spitzberg, R.M. and Richards, W. (1975) Broad band filters in the human visual system. Vis. Res. 15, 837-841.
- Stecher, S.; Sigel, C.; and Lange, R.V. (1973) Composite adaptation and spatial frequency interactions. Vis. Res. 13, 2527-2531.
- Stiles, W.S. and Burch, J.M. (1959) N.P.L. Colour-matching investigations: Final Report. Optica Acta 6, 1-26.
- Stromeyer, C.F., III and Julesz, B. (1972) Spatial frequency masking in vision: cortical bands and spread of masking. J. Opt. Soc. Am. 62, 1221-1231.
- Stultz, K.F. and Zweig, H.J. (1959) Relation between graininess and granularity for black-and-white samples with non-uniform granularity spectra. J. Opt. Soc. Am. 49, 693-702.
- Sullivan, G.D.; Georgeson, M.A. and Oatley, K. (1972) Channels for spatial frequency selections and the detection of single bars by the human visual system. Vis. Res. 12, 383-394.
- Tolhurst, D.J. (1972) Adaptation to square-wave gratings: inhibition between spatial frequency channels in the human visual system. J. Physiol. Lond. 226, 231-249.

- Triendl, E.E. (1973) Texturerkennung und texturreproduktion. Kybernetik 13, 1-5.
- Uozumi, J; Fujii, H. and Asakura, T. (1977) Further computer simulation study of image speckle patterns with relation to object surface profile. J. Opt. Soc. Am. 67, 808-814.
- Waltz, D. (1975) Understanding line drawings of scenes with shadows. pp. 19-92 in: The Psychology of Computer Vision. P.H. Winston (Ed.), McGraw-Hill: New York.
- Watanabe, A.; Mori, T.; Nagata, S. and Hiwatashi, K. (1968) Spatial sine-wave responses of the visual system. Vis. Res. 8, 1245-1263.
- Watson, A.B. and Nachmias, J. (1977) Patterns of temporal interaction in the detection of gratings. Vis. Res. 17, 893-902.
- Wilson, H.R. (1978) Personal communication.
- Wilson, H.R. and Giese, S.C. (1977) Threshold visibility of frequency gradient patterns. Vis. Res. 17, 1177-1190.
- Winston, P.H. (1975) The Psychology of Computer Vision. McGraw-Hill: New York.
- Wohlwill, J.F. (1962) The perspective illusion: Perceived size and distance in fields varying in suggested depth, in children and adults. J. Exp. Psychol. 64, 300-310.
- Wright, W.D. (1928) A redetermination of the trichromatic coefficients of the spectral colours. Trans. Opt. Soc. 30, 141.
- Wyszecki, G. and Stiles, W.S. (1967) Color Science. Wiley and Sons: New York.

Appendix I

Memory Reconfiguration Interface

(E. Black)

An important component of the visual display is the refresh memory, which allows us to store and display any arbitrary pattern (or picture). However, this memory is not always in use, and rather than sitting idle, it is desirable to make it available to the PDP 11 to expand core. For example, the PDP 11 has an address space of only 32K of 16 bit words. The "memory reconfiguration interface" is a scheme for making the large amount of 18 bit word refresh memory available to the 16 bit word processor when the refresh capability is not needed. This interface thus allows the refresh memory to be used in modes other than for video storage, greatly increasing its versatility and usefulness at little increase in cost.

Example Illustrating Interface

A common processor is the PDP 11 which has an address space of 32K ($K=1024$) of 16 bit words. Because the I/O device interfaces take up 4K of the memory, the available program memory is 28K. Customarily, however, the processor is purchased with considerable less memory, for example only 12K (the minimum unit is 4K). Thus there is usually a considerable gap--up to 24K--between the top of the memory supplied with the processor and the I/O device interfaces. Depending on the configuration of the interface, this gap may be filled by the refresh memory.

In the above example where the processor came with 12K, the remaining 16K gap may be filled as follows:

Address	Size, Description
Bottom of Memory	
0	12K processor memory (furnished)
12K	16K refresh memory (optional)
28K	4K I/O device interfaces (furnished)
Top of Memory	

In the above example, 16K of refresh memory completely fills the unused portion of processor memory. In the event that only 12K of additional processor memory were needed, this 12K gap could be only partially filled with the reconfigured refresh memory based on 8K pages. A different reconfiguration, such as one based on 4K pages of memory would be required to completely fill a 12K memory gap.

For refresh memory reconfigured with 8K pages for processor use, there will be eight such pages to give the total of 64K of video memory necessary to refresh a 440*440 element display having 6 bit brightness levels. When this same memory is used by the processor, an 8K page must be selected by programming some control bits in an I/O buffer. Once this specification is accomplished, the processor must view the memory in terms of its normal 16 bit word, rather than as the 18 bit words used for video refresh. Again, this change is controlled by the processor, which allows either 8 or 9 bit bytes to be read or written from a 16 bit processor word. In refresh mode, four 9 bit processor transactions are reconfigured to give six 6 bit video bytes (total 36 bits in each case, or two 18 bit words), whereas in processor mode the same 36 bit unit is used to create two 16 bit words with four bits remaining. These 4 remaining bits may still be used, but there is a software cost in accessing them.

In the 9 bit mode, four page-select bits are required to select one out of sixteen 8K pages. In extra processor memory mode, however, only three page-select bits are required

since there are half as many 18 bit words as 9 bit words. These reconfigurations are all accomplished by the interface so that regardless of the mode of use, the alterations in the memory are invisible to the user (i.e., the processor or refresh device).

Implementation

A) Specific Case: PDP 11/10 with 16K memory & Video refresh 64K by 18 bit words, 6 bit video.

The implementation at the reconfiguration interface is a fairly straightforward multiplexing of the address and data lines that control the core memory. For multiplexing data lines, the general scheme is to transfer half a processor word (8 bits) to and from half a memory word (9 bits). Thus two bits in the memory are ignored (these are the most significant bits of each half-word). When the processor needs to store or retrieve the full 9 bits of a memory half-word, it reconfigures the interface so that the 9 least-significant processor bits are transferred to and from a memory half-word. In this case the upper (most-significant) 7 bits in the processor word are ignored. The scheme may be illustrated by the following table:

The scheme may be illustrated by the following table:

DATA MULTIPLEXING					
Reading (to CPU):			Writing (from CPU):		
CPU bit	Processor mode	9-bit mode	Memory bit	Processor mode	9-bit mode
p0	from m0	or m0, 9	m0	from p0	p0
p1	m1	m1, 10	m1	p1	p1
p2	m2	m2, 11	m2	p2	p2
p3	m3	m3, 12	m3	p3	p3
p4	m4	m4, 13	m4	p4	p4
p5	m5	m5, 14	m5	p5	p5
p6	m6	m6, 15	m6	p6	p6
p7	m7	m7, 16	m7	p7	p7
p8	m9	m8, 17	m8	-	p8
p9	m10	-	m9	p8	p0
p10	m11	-	m10	p9	p1
p11	m12	-	m11	p10	p2
p12	m13	-	m12	p11	p3
p13	m14	-	m13	p12	p4
p14	m15	-	m14	p13	p5
p15	m16	-	m15	p14	p6
			m16	p15	p7
			m17	-	p8

Note: pi indicates a processor bit number, mi a memory bit number.

The figure is best viewed as two tables, one for reading, and one for writing. In each case the target device bit numbers are in the left-most column (CPU for reading, Memory for writing). Let us first consider the processor mode. Here we are concerned with the first two columns in each table. When writing, we see that processor bit 0 (p0) is written into memory bit 0, similarly, p1 is written into memory bit one, and so on up to p7. Nothing is stored in memory bit 8, since the processor has one bit less than the memory in each half-word. The second memory half-word has bits numbered 9-17, for a total of 9 again. The processor has bits numbered p8-p15, and once again, the lowest bit from the processor byte is paired with the lowest bit of the memory byte. Thus we see p8 is written into memory bit 9, etc., and nothing is written into bit 17.

In order to retrieve what was written, the read operation in this mode must simply reverse the bit mappings, and this is seen in the first two columns of the read table. Here the processor is the target (first column): memory bit 0 (m0) will be read into processor bit 0, and so on up to m7 and processor bit 7. Now the next most significant processor bit is bit 8, which was written into memory bit 9 because of the byte-length difference. Thus we see that on reading, m9 must be read into processor bit 8. The most significant processor bit is bit 15, which is seen to be read back from memory bit 16 (m16), and once again, m17 is ignored.

In the 9 bit mode, things are a little more tricky, as may be seen in the right-most column of each table. Consider writing: here, the processor is sending a 9 bit byte in its bits numbered p0-p8. These bits are written into either the high or low byte of memory (depending on an address bit). Therefore bits p0-p8 appear opposite both bytes, memory bits 0-8 and 9-17. When reading, the inverse map must be performed so that CPU bits 0-8 retrieve the high or low byte (depending on the same address bit). Thus both m0-m8 and m9-m17 appear as candidates to get read into the nine processor bits--the particular byte is selected by a multiplexor that is switched by the byte addressing bit in the reconfiguration interface.

ADDRESS MULTIPLEXING

Address multiplexing is necessary because of the conflicts which arise in addressing the EMM memory both normally and in a 9-bit byte mode. Because of the conflicts, the addresses which are generated by the CPU during an instruction must be interpreted differently by the memory, depending on the mode of addressing being used. The normal mode is described first:

The PDP 11 has a 32K address space. It takes sixteen bits to address such a space. The EMM memory was designed to fill one quarter of that space and so has fourteen bits of address available (two bits are required to specify one of four quarters). Since the EMM

memory requires seventeen bits to specify a byte uniquely, there remains a need to generate three more bits. These three remaining bits are written into an I/O buffer, and select one-eighth of the EMM memory (two raised to the third power is eight) until such time as the program changes them.

In 9 bit byte mode, the CPU cannot use its built-in byte operations, since one cannot assume that more than eight bits (a PDP 11 byte) will be presented to the memory. A word transaction (16 bits) must be used to specify a byte of memory, effectively reducing the number of address lines by one. An extra bit in the I/O register must therefore be used for this mode.

The address mappings are chosen to provide uniform access across modes--sequential bytes are retrieved in either processor or 9 bit byte mode in the order in which they would be displayed in refresh graphics mode. The following table illustrates the address multiplexing:

Address bits:

Memory address	Processor mode	9-bit mode	
ma 0	pa 0	pa 1	; byte selection
ma 1	pa 1	pa 2	; first word address bit
ma 2	pa 2	pa 3	; second word address bit
ma 3	pa 3	pa 4	
ma 4	pa 4	pa 5	
ma 5	pa 5	pa 6	
ma 6	pa 6	pa 7	
ma 7	pa 7	pa 8	
ma 8	pa 8	pa 9	
ma 9	pa 9	pa 10	
ma 10	pa 10	pa 11	
ma 11	pa 11	pa 12	
ma 12	pa 12	pa 13	
ma 13	pa 13	i/o bit 0	; 4K page selection
ma 14	i/o bit 0	i/o bit 1	; 8K page selection
ma 15	i/o bit 1	i/o bit 2	; 16K page selection
ma 16	i/o bit 2	i/o bit 3	; 32K page selection

Note: pa i indicates a CPU address (on the unibus)
ma i indicates a memory address bit.

B) General Case:

Although implementation of the reconfiguration interface is described and implemented in terms of a specific case, the scheme has generality to a wide variety of other applications. Whenever excess memory based upon words of bit length M is available, the same scheme can be used to reconfigure the memory for use by another system utilizing words of length $(M-n)$ bits.

Appendix II

Special Graphics System

In order to generate texture patterns from their spatial frequency (Fourier) components, we have designed and have built a special graphics display. This display allows us to generate a 440 x 440 x 6 bit brightness pattern consisting of complex (computer-generated) sinusoids whose contrast may be altered every 20 msec. Both the sums and products of up to six components may be displayed with variable contrast.

More specifically, the special visual display consists of 7 subsystems as follows:

1. Monitors: Conrac SNA 17/C (2)
Monochrome television monitors
2. Operator Controls: Two channels, each with independent control of three sinusoidal or other component amplitudes and the $a(x)*a(y)$ product term. Control boxes are on extension cables for convenience and flexibility of location. A six-channel A/D converter digitizes the the control settings for input to the computer.
3. Function Table Computer: A dedicated PDP 11/10 Minicomputer is used to monitor the operator controls and calculate $a(u)$ and $b(u)$ function tables in accordance with the operator control settings, where

$$a(u) = \sum_{i=1}^3 A_i \sin(2\pi f_i u + \phi_i)$$

$$b(u) = \sum_{i=1}^3 B_i \sin(2\pi f_i u + \phi_i)$$

4. Video Function Generators: Two identical custom designed video generators are provided to store the computed function tables and generate a video luminance signal of the form:

$$L_A(X,Y) = 1 + a(x) + a(y) + K_A a(x) a(y)$$

$$L_B(X,Y) = 1 + b(x) + b(y) + K_B b(x) b(y)$$

Provision is made for adding an external video signal.

5. Scan Generator: A custom-designed digital scan generator generates raster coordinates, synchronizing signals and control signals.

6. Video Refresh System: A custom-designed video refresh system is provided to allow an arbitrary two-dimensional pattern to be added to the texture display. The refresh system employs a standard core memory of 32,768 thirty-six bit words and can store 196,608 picture elements (pixels) with 6 bit (64 level) gray scale.

The EMM Micromemory 3000 series has been used for the core memory. Four 3000DD (16K x 18 bit) cards are mounted in a 5 1/4" high chassis together with a control and video output card, power supply and cooling fans. The control card circuit provides an alternate mode of operation in which four 108 x 108 checkerboard patterns can be stored and refreshed. The PDP 11/10 has control of mode selection and can select which of the four patterns is to be displayed.

The video refresh core memory can also be easily converted on site to a general purpose RAM. (See Appendix I.)

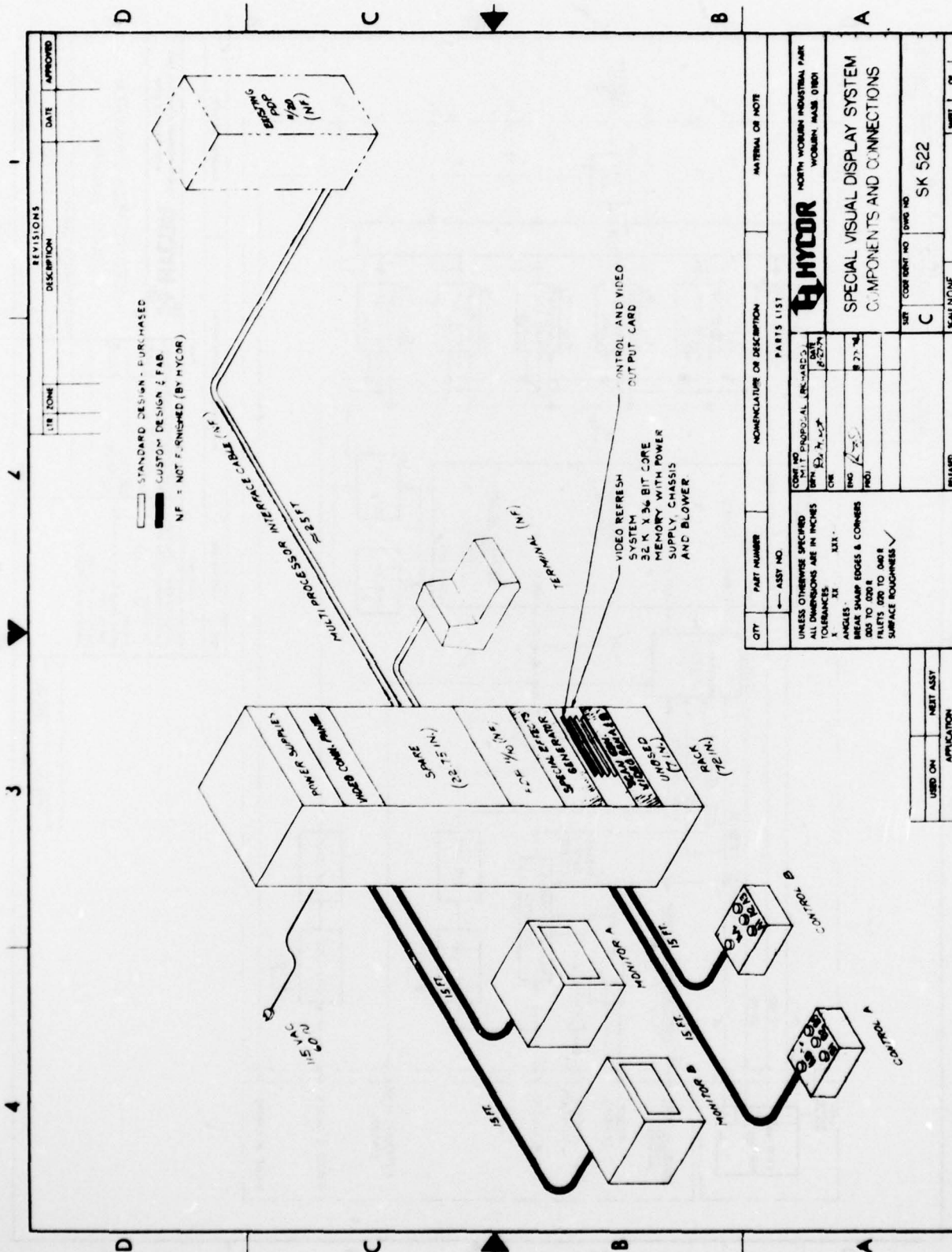
7. Video Interconnect Panel: A video interconnect panel is available to permit easy and flexible interconnection of video signals. The panel also contains eight adjustable DC voltage sources and a video integrator for use with the special

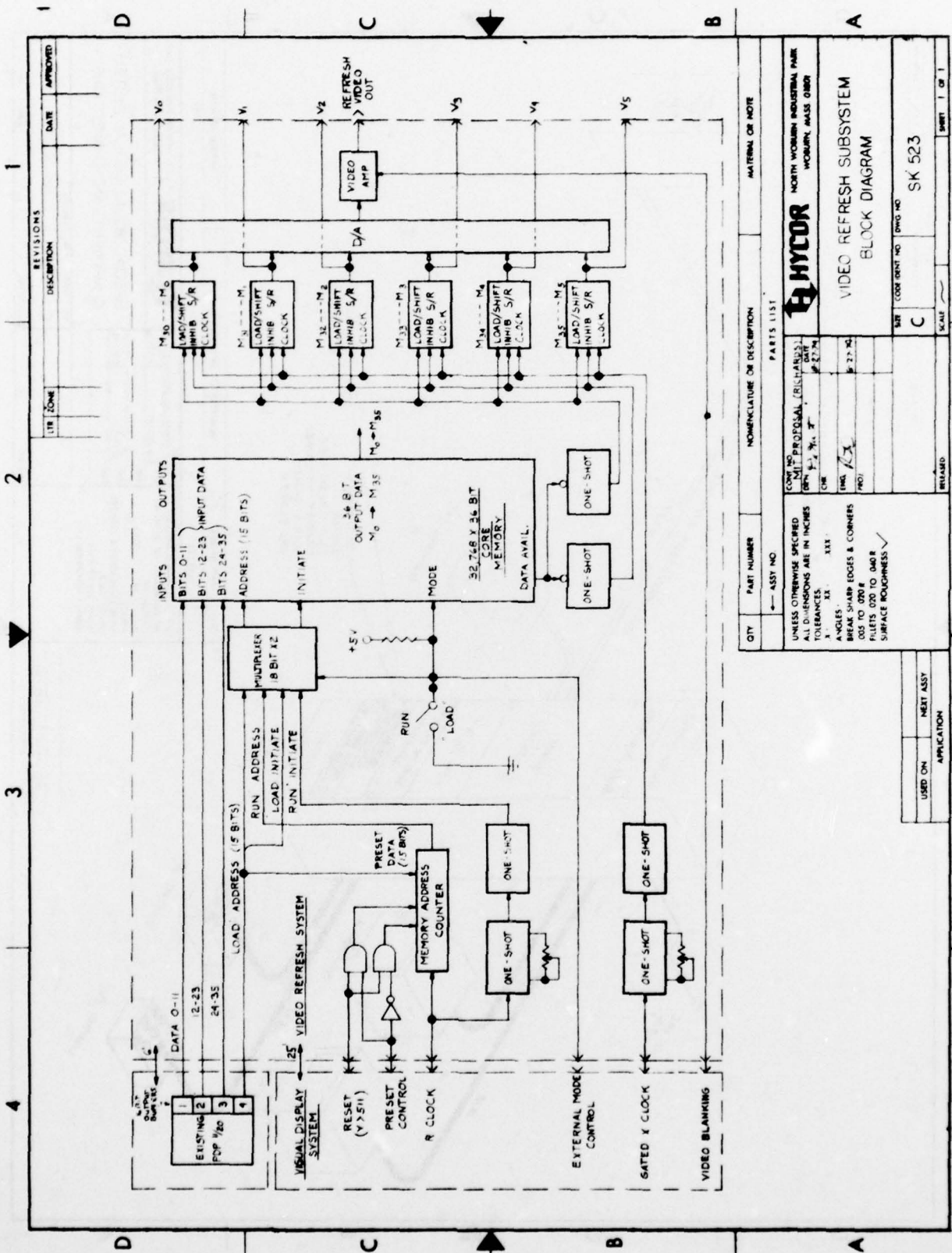
effects generator and video multiplexer. The prints describing these components in more detail are given in Appendix I.


Appendix III

System Schematics

Figure SK522:	Special Visual Display System Components and Connections
Figure SK523:	Video Refresh Subsystem Block Diagram
Figure SK524:	Scan Generator, Block Diagram
Figure SK525:	Video Generator, Block Diagram
Figure SK526:	Special Effects Generator
Figure SK527:	Special Visual Display System, Block Diagram





QTY	PART NUMBER	← ASSY NO	UNLESS OTHERWISE SPECIFIED ALL DIMENSIONS ARE IN INCHES TOLERANCES ARE: FRACTIONS DECIMALS ANGLES .XXX BREAK SHARP EDGES & CORNERS .005 TO .070 R FILLETS .020 TO .040 R SURFACE ROUGHNESS ✓
NOMENCLATURE OR DESCRIPTION		PARTS LIST	CONT NO MIT PROPOSAL DISPOSITION (REV. 1-2-74) (REV. 1-2-74) CHG INCL <i>1/2</i> (REV) DATE 12-27-82
MATERIAL OR NOTE		 NORTH WOODBURN INDUSTRIAL PARK WOODBURN, MASS 01890	DATE COORDINATE NO DWG NO C SK 523
		SCALE	SHEET 1 OF 1 RELEASED

